University of Alberta

ANN modeling of $PM_{2.5}$ for the cities of Calgary and Edmonton

by

Madhan Selvaraj

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Master of Science

in

Environmental Engineering

Department of Civil and Environmental Engineering

Edmonton, Alberta

Fall 2005

Library and Archives Canada

Published Heritage Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et Archives Canada

Direction du Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-09282-3
Our file Notre référence
ISBN: 0-494-09282-3

NOTICE:
The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

AVIS:
L'auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l'Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L'auteur conserve la propriété du droit d'auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n'y aura aucun contenu manquant.

Canada

ABSTRACT

Artificial neural networks (ANNs) were applied to model and forecast concentrations of particulate matter with aerodynamic diameter less than 2.5 μm ($PM_{2.5}$) in Edmonton and Calgary. $PM_{2.5}$ has been linked to respiratory illness and premature mortality in numerous recent epidemiology studies, and is a growing concern for public health protection. This research examines the potential for using ANN models to assess $PM_{2.5}$ control measures, to provide an additional source of $PM_{2.5}$ information, and to predict situations that may threaten the health of the public.

Initial results for the $PM_{2.5}$ models were poor. When the data were subdivided into seasons and individual ANN models developed for each season, prediction performance and modelling capability increased substantially for most seasons. For the Edmonton virtual monitor model, the best model was for the spring and winter seasons, with a coefficient of multiple determinations (R^2) of 0.72 when applied to the production data set. This was followed by fall ($R^2 = 0.66$), and summer ($R^2 = 0.41$). The models for the winter, fall, and summer forecasts showed performance improvements, with R^2 values of 0.78, 0.76, and 0.58 respectively. The spring forecast model also performed well at $R^2 = 0.65$.

For Calgary, the spring and fall virtual monitor models had R^2 values of 0.61 and 0.72 respectively. The summer virtual monitor model performance was at $R^2 = 0.71$, followed by winter at $R^2 = 0.46$. The fall had a R^2 of 0.68, followed by summer ($R^2 = 0.71$), spring ($R^2 = 0.67$), and then winter ($R^2 = 0.38$) for forecast model.

Forest fires near Edmonton and wet deposition effects may affect the summer data for Edmonton, limiting the success of these models. In Calgary, the ANN models had difficulty modelling winter data, possibly due to the erratic winters in the area. Improvements to the performance of the ANN models should be investigated in future studies. Inclusion of more high $PM_{2.5}$ events in the training data and time series effects are possible refinements to the models.

ACKNOWLEDGEMENTS

A special thanks to Diane Su and Kevin McCullum for their guidance in various stages of this research. Also, thanks to Dr. Warren Kindzierski and Brian Wiens for their insights into the theoretical aspects of this research.

I also liked to thank my brother and parents without them I would not have pursued my masters study in Canada.

Finally, I am indebted to my supervisor Dr. Gamal El-Din for his guidance and advice during my entire research work.

TABLE OF CONTENTS

1.0 GENERAL INTRODUCTION 1

2.0 PARTICULATE MATTER $_{2.5}$ FORMATION AND THEIR EFFECTS 3

2.1 Introduction 3

2.2 Particulate matter- background 3

2.2.1 Sources of $PM_{2.5}$ 4

2.2.2 Sinks for particles 7

2.3 Role of meteorology in $PM_{2.5}$ formation and dispersion 7

2.3.1 Theory 7

2.3.2 Meteorological studies related to PM2.5 8

2.4 Health effects 8

2.5 Material and vegetative effects 9

2.6 Visibility effects 9

2.7 Standards for $PM_{2.5}$ 10

2.8 Conclusions 11

2.9 References 11

3.0 RECEPTOR MODELING TECHNIQUES AND ANN 16

3.1 Introduction 16

3.2 Receptor modelling techniques 16

3.3 Artificial Neural Networks – A theoretical background 18

3.4 ANNs and air pollution 21

3.5 Conclusions 23

3.6 References 23

4.0 EDMONTON EAST $PM_{2.5}$ MODELS 28

4.1 Introduction 28

4.2 Artificial Neural Networks 29

4.3 Data and methods 30

4.3.1 Study Area 30

4.3.2 Model Inputs 30

4.3.3 Model Development 33

4.3.4 Systematic Approach 33

4.4 Results and discussion 35
4.4.1 Virtual Monitor 36
4.4.1.1 Spring model 37
4.4.1.2 Summer season model 38
4.4.1.3 Fall season 40
4.4.1.4 Winter season 41
4.4.2 Important input variables 43
4.4.3 Forecast models 45
4.4.3.1 Spring season 45
4.4.3.2 Summer season 45
4.4.3.3 Fall season 47
4.4.3.4 Winter season 48
4.4.4 Check for persistence 49
Virtual monitor 50
Forecast model. 52
4.5 Conclusions 55
4.6 References 56
5.0 CALGARY EAST $PM_{2.5}$ MODELS 61
5.1 Introduction 61
5.2 Data and Methods 62
5.2.1 Study area 62
5.2.2 Artificial Neural Network 62
5.2.3 Methodology 64
5.3 Results and Discussion 66
5.3.1 Prediction model 66
5.3.1.1 Spring model 67
5.3.1.2 Summer model 68
5.3.1.3 Fall model 70
5.3.1.4 Winter model 71
5.3.2 Prediction model with lagged data as an input 73
5.3.2.1 Spring model 73
5.3.2.2 Summer model 74

5.3.2.3 Fall season 74
5.3.2.4 Winter season 75
5.3.3 Forecast model 76
5.3.3.1 Spring model 77
5.3.3.2 Summer model 77
5.3.3.3 Fall model 78
5.3.3.4 Winter model 79
5.3.4 Forecast model with lagged data as an input 80
5.3.4.1 Spring model 80
5.3.4.2 Summer model 81
5.3.4.3 Fall model 81
5.3.4.4 Winter model 82
5.3.5 Important Input variables 84
5.3.5.1 Spring season 84
5.3.5.2 Summer season 85
5.3.5.3 Fall season 86
5.3.5.4 Winter season 87
5.4 Conclusions 88
5.5 References 89
6.0 FINDINGS AND RECOMMENDATIONS 93
APPENDIX-A 95
APPENDIX-B 104

LIST OF TABLES

Table 2- 1 Emission amount from various sources for $PM_{2.5}$ in Alberta for the year 2000..... 5

Table 2- 2 Visibility condition prevalent in different Canadian location (McDonald and Shepherd, 2004). 10

Table 2- 3 Standards for $PM_{2.5}$ followed in various countries 10

Table 4- 1 Number of hidden layer neurons and training epochs setting for evaluating hidden and output layer activation function..... 35

Table 4- 2 Basic statistics for different season historical data. 36

Table 4- 3 Summary of architecture followed in developing seasonal virtual monitor models. 37

Table 4- 4 Summary of architecture followed in developing seasonal virtual monitor models. 43

Table 4- 5 Change in R^2 value with respect to removal of particular variable 44

Table 4- 6 Input variables important in the prediction of $PM_{2.5}$ for all the seasons... 46

Table 4- 7 Summary of architecture followed in developing seasonal forecast models. 50

Table 4- 8. Summary of virtual monitor and forecast models with lagged data as an input. 55

Table 5- 1 Parameters used in developing models for different seasons. 65

Table 5- 2 Number of hidden layer neurons and training epochs settings for evaluating hidden and output layer activation functions. 66

Table 5- 3 Type of activation function that is used in model development for different seasons based on previous study (Su et al., 2004). 67

Table 5- 4 Summary of virtual monitor model characteristics with and without lagged data as an input..... 76

Table 5- 5 Summary of forecast model characteristics with and without lagged data as an input. 83

Table 5- 6 Change in R^2 value with respect to removal of particular variable 87

LIST OF FIGURES

Figure 2-1 Formation of particulate matter from various sources 4
Figure 2-2 Formation of particles from carbonaceous compounds. 6
Figure 3- 1 Source-receptor relationship 17
Figure 3- 2 A typical Neural Networks 19
Figure 3- 3 Internal working of a neuron 20
Figure 4- 1 Optimization of model architecture for spring season virtual monitor model 38
Figure 4- 2 Edmonton East spring virtual monitor model. $R^2 = 0.65$ 39
Figure 4- 3 Optimization of model architecture for summer season. 39
Figure 4- 4 Edmonton East summer virtual monitor model. $R^2 = 0.41$ 40
Figure 4- 5 Optimization of model architecture for fall season 41
Figure 4- 6 Edmonton East fall virtual monitor model. $R^2 = 0.72$. 41
Figure 4- 7 Optimization of model architecture for winter season 42
Figure 4- 8 Edmonton East winter virtual monitor model. $R^2 = 0.72$. 42
Figure 4- 9 Edmonton East spring forecast model. $R^2 = 0.72$. 47
Figure 4- 10 Edmonton East summer forecast model. $R^2 = 0.58$ 47
Figure 4- 11 Edmonton fall forecast model. $R^2 = 0.76$ 48
Figure 4- 12 Edmonton winter forecast model. $R^2 = 0.78$ 49
Figure 4- 13 Edmonton East Spring virtual monitor model with lagged data. $R^2 = 0.80$. 51
Figure 4- 14 Edmonton East summer virtual monitor model with lagged data. $R^2 = 0.65$. 51
Figure 4- 15 Edmonton East fall virtual monitor model with lagged data. $R^2 = 0.80$. ... 52
Figure 4- 16 Edmonton East winter virtual monitor model with lagged data. $R^2 = 0.84$. 52
Figure 4- 17 Edmonton East Spring forecast model with lagged data. $R^2 = 0.78$. 53
Figure 4- 18 Edmonton East summer virtual monitor model with lagged data. $R^2 = 0.58$. 53
Figure 4- 19 Edmonton East Fall forecast model with lagged data. $R^2 = 0.79$. 54

Figure 4- 20 Edmonton East Winter forecast model with lagged data. $R^2 = 0.74$. 54
Figure 5- 1 A three-layered neural network 63
Figure 5- 2 Optimization of model architecture for spring season virtual monitor model..... 68
Figure 5- 3 Calgary East spring prediction model. $R^2 = 0.61$..... 68
Figure 5- 4 Optimization of model architecture for summer season prediction model... 69
Figure 5- 5 Calgary East summer prediction model. $R^2 = 0.71$..... 69
Figure 5- 6 Optimization of model architecture for fall season prediction model. 70
Figure 5- 7 Calgary East Fall prediction model. $R^2 = 0.72$. 71
Figure 5- 8 Optimization of model architecture for winter season prediction model. 72
Figure 5- 9 Calgary East winter prediction model. $R^2 = 0.46$. 72
Figure 5- 10 Calgary East Spring prediction model with lagged data. $R^2 = 0.76$. 73
Figure 5- 11 Calgary East Summer prediction model with lagged data. $R^2 = 0.74$..... 74
Figure 5- 12 Calgary East fall prediction model with lagged data. $R^2 = 0.85$..... 75
Figure 5- 13 Calgary East winter prediction model with lagged data. $R^2 = 0.50$..... 76
Figure 5- 14 Calgary East spring forecast model with lagged data. $R^2 = 0.67$..... 77
Figure 5- 15 Calgary East summer forecast model. $R^2 = 0.68$. 78
Figure 5- 16 Calgary East fall forecast model $R^2 = 0.71$..... 79
Figure 5- 17 Calgary East winter forecast model $R^2 = 0.46$..... 80
Figure 5- 18 Calgary East spring forecast model with lagged data as an input, $R^2 = 0.63$. 81
Figure 5- 19 Calgary East summer forecast model with lagged data as an input, $R^2 = 0.78$..... 82
Figure 5- 20 Calgary East fall forecast model with lagged data as an input, $R^2 = 0.78$.... 82
Figure 5- 21 Calgary East winter forecast model with lagged data as an input $R^2 = 0.37$. 83
Figure 5- 22 Change in R^2 value due to removal of input variable for spring season model..... 85
Figure 5- 23 Change in R^2 value due to removal of input variable for summer season model..... 86

Figure 5- 24 Change in R^2 value due to removal of input variable for fall season model 86
Figure 5- 25 Change in R^2 value due to removal of input variable for winter season model..... 87
Figure A- 1 Network vs. Actual $PM_{2.5}$ for Edmonton East spring virtual monitor model..... 96
Figure A- 2 Edmonton East spring virtual monitor residual analysis 96
Figure A- 3 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model..... 97
Figure A- 4 Edmonton East summer virtual monitor residual analysis 97
Figure A- 5 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model..... 98
Figure A- 6 Edmonton East fall virtual monitor residual analysis..... 98
Figure A- 7 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model..... 99
Figure A- 8 Edmonton East winter virtual monitor residual analysis 99
Figure A- 9 Network vs. Actual $PM_{2.5}$ for Edmonton East spring forecast model.... 100
Figure A- 10 Edmonton East spring forecast residual analysis..... 100
Figure A- 11 Network vs. Actual $PM_{2.5}$ for Edmonton East summer forecast model. 101
Figure A- 12 Edmonton East summer forecast residual analysis..... 101
Figure A- 13 Network vs. Actual $PM_{2.5}$ for Edmonton East fall forecast model. 102
Figure A- 14 Edmonton East fall forecast residual analysis 102
Figure A- 15 Network vs. Actual $PM_{2.5}$ for Edmonton East winter forecast model. ... 103
Figure A- 16 Edmonton East winter forecast residual analysis..... 103
Figure B- 1 Network vs. Actual $PM_{2.5}$ for Calgary spring prediction model..... 105
Figure B- 2 Calgary East spring prediction model residual analysis. 105
Figure B- 3 Network vs. Actual $PM_{2.5}$ for Calgary summer prediction model..... 106
Figure B- 4 Calgary East summer prediction model residual analysis. 106
Figure B- 5 Network vs. Actual $PM_{2.5}$ for Calgary fall prediction model 107
Figure B- 6 Calgary East fall prediction model residual analysis..... 107
Figure B- 7 Network vs. Actual $PM_{2.5}$ for Calgary winter prediction model. 108
Figure B- 8 Calgary East winter prediction model residual analysis..... 108

Figure B- 9 Network vs. Actual $PM_{2.5}$ for Calgary spring prediction model with lagged data as an input..... 109
Figure B- 10 Calgary East spring prediction model (lagged data) residual analysis. .. 109
Figure B- 11 Network vs. Actual $PM_{2.5}$ for Calgary summer prediction model with lagged data as an input..... 110
Figure B- 12 Calgary East summer prediction model (lagged data) residual analysis. 110
Figure B- 13 Network vs. Actual $PM_{2.5}$ for Calgary fall prediction model with lagged data as an input..... 111
Figure B- 14 Calgary East fall prediction model (lagged data) residual analysis..... 111
Figure B- 15 Network vs. Actual $PM_{2.5}$ for Calgary winter prediction model with lagged data as an input..... 112
Figure B- 16 Calgary East winter prediction model (lagged data) residual analysis... 112
Figure B- 17 Network vs. Actual $PM_{2.5}$ for Calgary spring forecast model..... 113
Figure B- 18 Calgary East spring forecast model residual analysis..... 113
Figure B- 19 Network vs. Actual $PM_{2.5}$ for Calgary East summer forecast model..... 114
Figure B- 20 Calgary East summer forecast model residual analysis..... 114
Figure B- 21 Network vs. Actual $PM_{2.5}$ for Calgary East fall forecast model. 115
Figure B- 22 Calgary East fall forecast model residual analysis. 115
Figure B- 23 Network vs. Actual $PM_{2.5}$ for Calgary East winter forecast model. 116
Figure B- 24 Calgary East winter forecast model residual analysis..... 116
Figure B- 25 Network vs. Actual $PM_{2.5}$ for Calgary East spring forecast model with lagged data as an input..... 117
Figure B- 26 Calgary East spring forecast model (lagged data) residual analysis..... 117
Figure B- 27 Network vs. Actual $PM_{2.5}$ for Calgary East summer forecast model with lagged data as an input..... 118
Figure B- 28 Calgary East summer forecast model (lagged data) residual analysis.... 118
Figure B- 29 Network vs. Actual $PM_{2.5}$ for Calgary East fall forecast model with lagged data as an input..... 119
Figure B- 30 Calgary East fall forecast model (lagged data) residual analysis 119
Figure B- 31 Network vs. Actual $PM_{2.5}$ for Calgary East winter forecast model with lagged data as an input..... 120

Figure B- 32 Calgary East winter forecast model (lagged data) residual analysis..... 120

GLOSSARY OF TERMS AND ABBREVIATIONS

CEPA	Canadian Environmental Protection Agency
CO	Carbon monoxide concentration
DAY	Input variable representing the day of the week.
DEV	Standard deviation of wind direction
EC	Environment Canada
EPA	Environmental Protection Agency
HOUR	Input variable representing the hour of the day.
MH	Mixing height
MTH	Input variable representing the month.
OPA	Opacity.
NH_4	Ammonium
NO	Nitric oxide concentration
NO_2	Nitrogen dioxide concentration
NO_x	Oxides of nitrogen concentration, the sum of NO and NO_2, in parts per million.
OC	Organic carbon
O_3	Ground-level ozone concentration, in parts per million.
$PM_{2.5}$	Particulate matter of diameter 2.5, in $\mu g/m^3$
PM_{10}	Particulate matter of diameter 10, in $\mu g/m^3$
ppm	Parts per million by volume.
RH	Relative humidity, %.
SO_2	Sulphur dioxide concentration, in parts per million.
t	One time period, generally in hours.
TEMP	Ambient temperature, in °C.
THC	Total hydrocarbons concentration, in parts per million.
WDR	Direction from which wind is blowing, in degrees from north.
WSP	Horizontal windspeed, in km/h.

1.0 GENERAL INTRODUCTION

The atmosphere as a sink was able to remove air pollutants, until modernization and rapid industrialization of the world. Increase in pollutants concentration above background levels has also led to a saturation of natural atmospheric dispersion behaviour. Increase in pollutant levels has also led to major episodes such as photochemical smog, acid at the local scale, and warming at the global scale.

Most pollution problems are intermingled. For example, a reduction in sulphur dioxide (SO_2) has shown to increase the nitrate particles because of the ammonium associated with the sulphate particle is available for ammonium nitrate. A decrease in VOC concentration has shown to increase the sulphate and nitrate concentration because more oxidants become available for oxidation of SO_2 to sulphate and NO_2 to nitrate. This clearly shows that air pollution is a complex phenomenon and all pollutants are intermingled. Receptors modeling techniques come in handy in these kinds of situations, since receptor modeling can be used to asses the present and future situations.

Particular matter 2.5 ($PM_{2.5}$), a potentially toxic compound consists of a "cocktail" of chemical compounds. $PM_{2.5}$ has wide range of sources. Because of its complex non-linear relationship with meteorology, understanding its formation is a great challenge for researchers.

Artificial neural networks (ANNs), an artificial Intelligence technique, is similar to the working of human brain and is a kind of receptor modelling technique, which can be used to model complex-non linear system. Because, of its previous success in various fields in modeling complex-non linear systems, ANNs was used in this project.

Two ANN models were developed in this project, each for Calgary and Edmonton. The two models developed were: a virtual monitor model that estimates $PM_{2.5}$ concentrations and a forecast model that predicts $PM_{2.5}$ concentrations in advance. The intent of the virtual monitor is to provide an alternate source of real-time $PM_{2.5}$ data when existing

monitors fail. This model may also be used to evaluate pollution reduction strategies and to test different ambient air scenarios. The forecast model provides a tool for predicting ambient $PM_{2.5}$ conditions that may harm human health. In these cases, advance knowledge of potentially poor air quality events would allow opportunity for the public to take steps to avoid excessive exposures.

2.0 PARTICULATE MATTER $_{2.5}$ FORMATION AND THEIR EFFECTS

2.1 Introduction

Particulate matter is a ubiquitous pollutant, reflecting the fact it has both natural and anthropogenic sources (Environment Canada, 1999 and 2000). $PM_{2.5}$ is onerous on human health because of their small size. $PM_{2.5}$ can have severe environmental and health effects. Research study on $PM_{2.5}$ in Canada has gained importance, because of the Canada wide standard (CWS), 30$\mu g/m^3$ (24-hr average) which is going to be implemented in the year 2010 (CCME, 2000).

2.2 Particulate matter- background

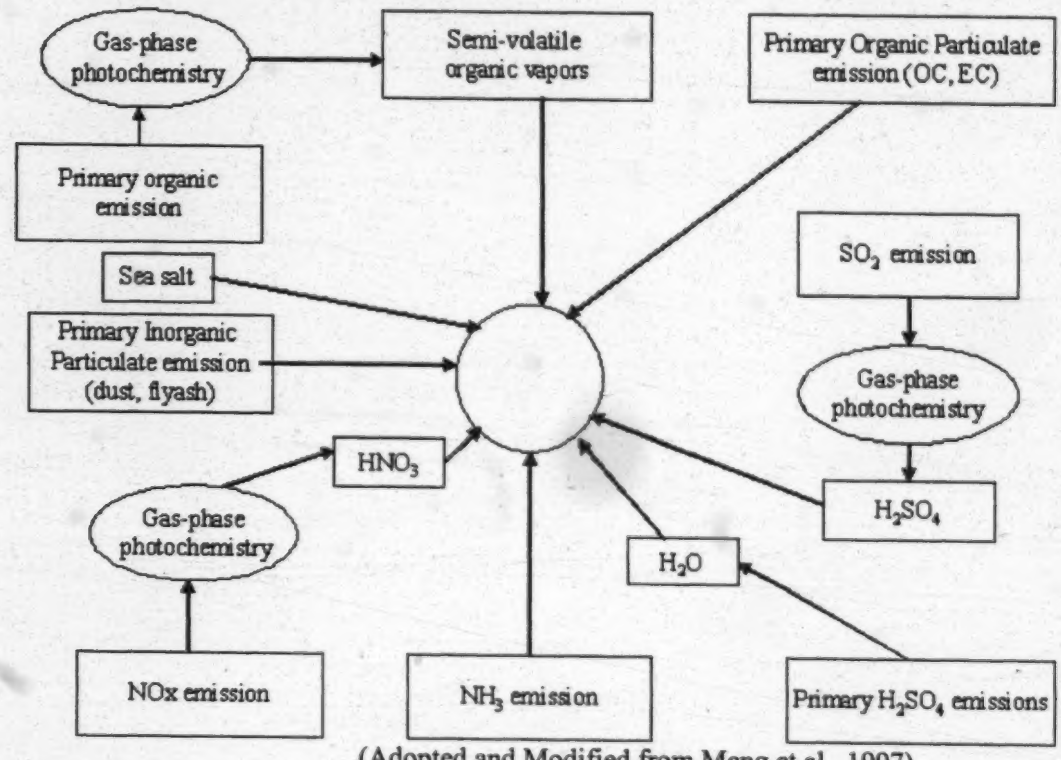
Particulate matter (PM) is composed of "cocktail" of constituents and chemical compounds. Some of the constituents are emitted directly in to the atmosphere (dust, soot, flyash, sea salt, etc.) whereas PM compounds are formed in the atmosphere through gas phase conversion (sulfate, nitrate, ammonium, organic compounds, etc.) (Seigneur, 2005; Pandis, 1995). Figure 2-1 shows the formation of PM. The word aerosol and PM is interchangeable in the field of air pollution (Seinfeld, 2004). Airborne particles can change their size and composition through various transformation processes such as condensation, coagulation, homogenous nucleation and chemical reactions. PM greater than 10μm in diameter are called as PM_{10}. $PM_{2.5}$ consists particles of diameter 2.5μm or less. Meanwhile, particles less than 1μm in size are referred to as ultra-fine particles. In this research, $PM_{2.5}$ is the particulate matter that is under study.

The principal components of $PM_{2.5}$ are high molecular weight organic compounds like alkanes, alkenes, and benzene, sulphates (ammonium sulphate, sulphuric acid, sulphur dioxide), nitrates, acid condensates, trace metals (lead, nickel, cadmium, chromium, and arsenic), and soot (CEPA/FPAC WGAQOG 1999; Dockery et al 1993; Perez et al. 2000). $PM_{2.5}$ levels are seasonal, observed to be highest in the typically long, cold winters of Alberta, when mixing height is lowest (Sandhu 1998).

2.2.1 *Sources of $PM_{2.5}$*

Common sources of $PM_{2.5}$ include vehicle emissions, processes involving combustion of fossil fuels, industry (Bates et al. 1990; Dockery et al. 1993; Viotti et al. 2002), wood burning, and wildfires (CEPA/FPAC WGAQOG 1999). In Calgary and Edmonton, additional sources of $PM_{2.5}$ are asphalt roofing manufacturing plants, atmospheric chemical transformation of sulphur dioxide and oxides of nitrogen, cement distribution centres, steel foundries, and forest fires. In Edmonton, sources also include petroleum refineries, cement kilns, and coal-fired power plants (Cheng et al. 1998).

In Alberta, dust from unpaved roads, forest fires, wood industry, agriculture (animals), prescribed burning, electrical power generation, dust from paved roads, residential wood combustion, off road diesel use, and heavy duty diesel vehicles are the major sources of $PM_{2.5}$ (Environment Canada, 1995).



(Adopted and Modified from Meng et al., 1997)

Figure 2-1 Formation of particulate matter from various sources

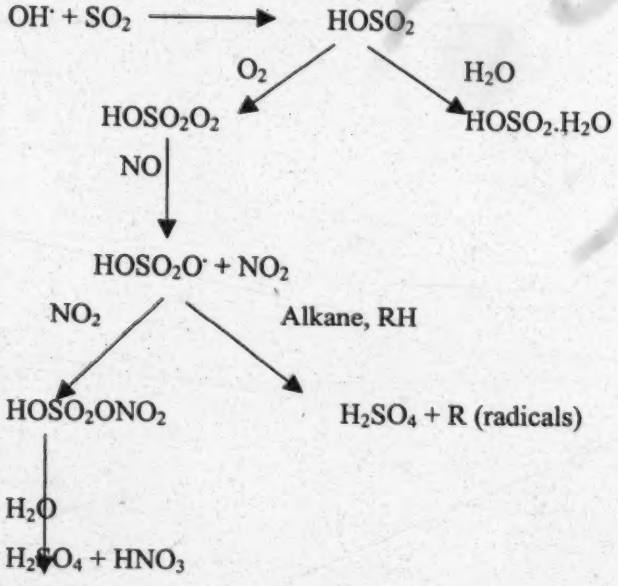
Table2-1 shows the $PM_{2.5}$ source emission in Alberta for the year 2000 (Environment Canada, 2000). In Canada, Ontario, Quebec had the highest pollution from all the sources.

Table 2- 1 Emission amount from various sources for $PM_{2.5}$ in Alberta for the year 2000 (Source:).

Sources	Emission (tonnes)
Industrial sources	18,628
Non-Industrial fuel combustion	7,623
Transportation	10,610
Incineration	155
Miscellaneous	1,142

Secondary particle formation is one important source of $PM_{2.5}$ formation. Formation of secondary particle requires the presence of chemicals such as SO_2, NH_3, NO_2, water, hydrocarbons as well as sunlight and ozone (ARC, 1998). Following are the reactions that contribute secondary particle formation in the atmosphere.

Sulphate particulate: sulphuric acid is the main constituent in the formation of sulphate-associated particulate formation (Seinfeld, 1980). Sulphuric acid is formed in different ways:



$H_2SO_4 + NH_3 \longrightarrow (NH_4)_2SO4$ (ammonium sulfate)

Sulfate salts are formed due to sulphuric acid reaction with ammonia in the atmosphere.

Nitrate particle: Nitrate-associated particulate formation is through a series of chemical reaction shown below:

$NO_2 + O_3 \longrightarrow NO_3 + O_2$

$NO_2 + NO_3 \longrightarrow N_2O_5$

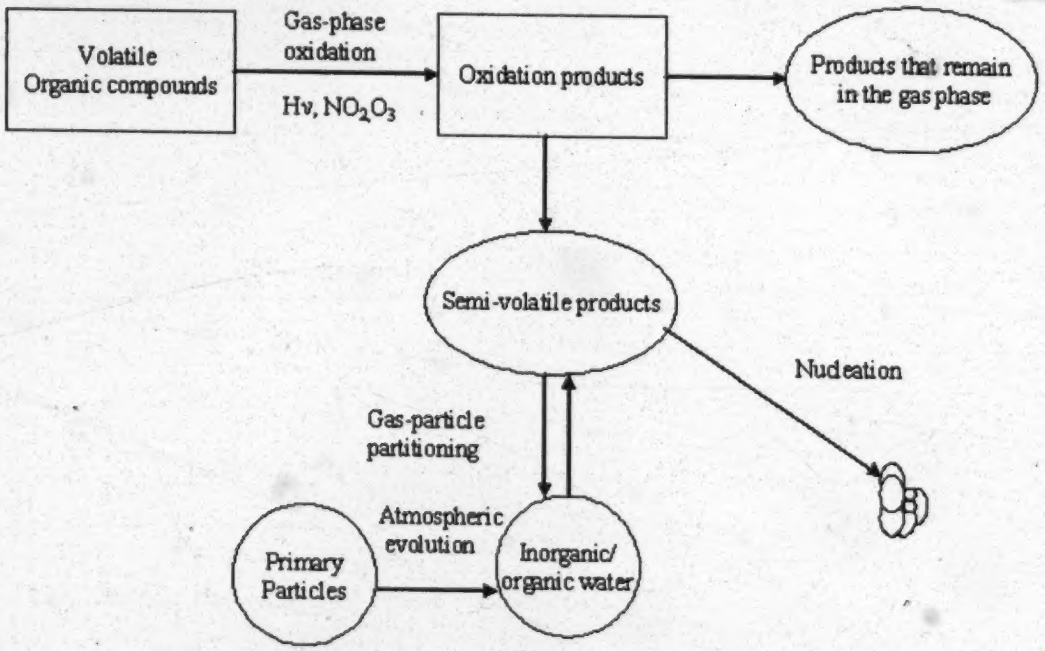
$N_2O_5 + H_2O \longrightarrow 2HNO_3$

Nitric acid reacts with ammonia to form ammonium nitrate.

$HNO_3 \longrightarrow NH_4NO_3$

Carbonaceous-associated particulate :

Organic carbon (OC) is the carbonaceous compound that is responsible for the secondary particulate formation (Figure 2-2). Secondary particle formation is through the oxidation of volatile organic compounds to a higher polarity and lower volatile reaction compounds (Seinfeld and Pandis, 1998).



(Adopted and Modified from Seinfeld and Pankow, 2003)

Figure 2-2 Formation of particles from carbonaceous compounds.

2.2.2 Sinks for particles

Particles are removed in the atmosphere by two major mechanisms: deposition at the earth surface, so called dry deposition and scavenging by water droplets through precipitation, wet deposition. The life time of PM ranges from few days to months based on their size. The life time of $PM_{2.5}$ is in days, with wet deposition playing a major role in the removal process (Pandis, 1995). Meanwhile, particles $\leq 0.1\mu m$ diameter have a short life span because of their high deposition velocities and coagulation with other particles. However, particles larger than few micrometers have greater settling velocity and subsequently have shorter life span.

2.3 Role of meteorology in $PM_{2.5}$ formation and dispersion

2.3.1 Theory

Stability of the atmosphere plays a vital role in the dispersion of any pollutant. Substances emitted in the atmosphere are gradually dispersed horizontally and vertically through the action of turbulence, and finally get mixed if sufficient time is available. This is called the mixing layer or planetary boundary layer. Turbulence may be convective and mechanical. Convective or thermal turbulence is dominant after sunup. Convective turbulence, accompanied by a strong wind speed is an ideal condition for dispersion of pollutants. The stability in this case is neutral. Mechanical turbulence is dominant in cloudy and night conditions.

Rate of change of temperature with respect to height is called lapse rate. If warm dry air is forced up in atmosphere, adiabatic cooling results in a temperature decrease of -10°C/km. This is called the dry adiabatic lapse rate. If actual lapse rate is greater than the dry adiabatic lapse rate, then the atmospheric condition is termed unstable and there is good dispersion of pollutants for elevated sources only. Atmospheric dispersion of pollutants is good if the lapse rate is equal to the dry adiabatic lapse rate. Stable condition prevails when the lapse rate is less than the adiabatic lapse rate. Increase in temperature with height is called an inversion and this prevalent in night and in the winter season. Stable atmospheric conditions prevail during temperature inversions, which act to trap air

pollutants. Calm wind speed, temperature inversions are some of the reasons for increased pollutant levels during winter season.

2.3.2 Meteorological studies related to $PM_{2.5}$

According to Laasko et al., (2003), particulate mass was found to vary diurnally and seasonally. In another study done in Vietnam (Hien et al., 2002), particulate concentration was found to be less in the monsoon season (June-august) compared to the winter season. Pathak et al., (2003) found that mixing height was correlated with sulphate, nitrate and ammonium concentrations in Hong Kong. Moreover, they found increase in sulphate and ammonium concentration due to long-range transport of aerosols. Even previous studies (Song and Carmichael, 1999) done on $PM_{2.5}$ found that long-range transport of aerosols had an influence on their composition. There was a case of long-range transport of dust influencing the $PM_{2.5}$ concentration (Fang et al., 1999; McKendry et al., 2001). Thus Meteorology plays a vital role in $PM_{2.5}$ formation and their distribution.

2.4 Health effects

PM has been linked to various respiratory and cardiovascular problems. Some of the common health effects related to PM are premature mortality, aggravation of respiratory and cardiovascular disease, aggravated asthma, and chronic bronchitis.
Creason et al. (2001), showed that exposure to $PM_{2.5}$ showed increase in heart rate variability in elderly people. Moreover, previous studies in the field of cardiology has also shown increase in pulse rate (peters et al., 1999; Pope et al., 1999a) and heart-rate variability (Pope et al., 1999b), when exposed to $PM_{2.5}$.

Exposure to fine particulate matter has shown increase in hospital admission and premature mortality (Dockery and Pope, 1994; Lebowitz, 1996; Lippmann and Thurston, 1996). Meanwhile, epidemiological studies have shown that air borne particulate matter is the reason for increased mortality rates in urban areas (Schwarz et al., 1996; Brunekreef et al., 1995; Xu et al., 1995, Dockery et al., 1994; Dockery et al., 1993).

2.5 Material and vegetative effects

The effects of PM on materials have been investigated in a limited way for metals, wood, stone, painted surfaces, electronics and fabrics. The deposition of PM on these materials may cause soiling and discoloration, thus reducing their aesthetic appeal. Exposure to PM also causes physical and chemical degradation of materials, through the action of acidic particles for example. The most obvious effect of particle deposition on vegetation is a physical smothering of the leaf surface. This will reduce light transmission to the plant, in turn causing a decrease in photosynthesis. Stomata (microscopic openings on the leaf surface) are also susceptible to blockage by particles. Particle composition may also be relevant, causing both direct chemical effects on the plant and indirect effects through impacts on the soil environment. Particle accumulation on the leaf surface may also increase the plant's susceptibility to disease.

2.6 Visibility effects

Visibility is the public perception about air quality (McDonald and Shepherd, 2004). Particles suspended in air interfere with the transmission of light through the atmosphere. Moreover, scattering and absorption of light by particles decreases the visibility conditions. Particulate associated sulphate concentration has shown to be the cause for visibility reduction ins Eastern United states (Seinfeld, 2004) compared to the Western United States where, sulphates and organics were the main contributor. In Canada, Ester, Alberta had the best visibility conditions for the year 1994-2000 (MacDonald and Shepherd, 2004). Meanwhile, the Windsor-Quebec City corridor had the worst visibility condition in Canada. Table 2-2 shows the visibility conditions prevalent in Canada.

2.7 Standards for $PM_{2.5}$

Adverse effects have been reported on the human and the environment by the pollutant $PM_{2.5}$. Based on this rationale, the Canada wide standard (CWS) for $PM_{2.5}$ was set as $30\mu g/m^3$ (CCME, 2000). Achievement to be based on the 98th percentile ambient measurement annually, averaged over 3 consecutive years. Table 2-3 shows the standards followed in other countries and in Canada.

Table 2- 2 Visibility condition prevalent in different Canadian location (McDonald and Shepherd, 2004).

Site	Visibility (Km)
Edmonton	100
Calgary	98
Ester	150
Toronto	51
Vancouver	83
Halifax	77
Winnipeg	97
Montreal	38
Ottawa	92
Hamilton	58

Table 2- 3 Standards for $PM_{2.5}$ followed in various countries

Location	Standards
Canada	$30\mu g/m^3$ (24-hr average)
United states	$65\mu g/m^3$ (24-hr average) $15\mu g/m^3$ (1-hr average)
Newfoundland, Canada	$25\mu g/m^3$ (24-hr average)
European Union	Follows USEPA standard

In Canada, 69 sites are maintained by Environment Canada (Environment Canada, 2004) for monitoring $PM_{2.5}$ and ozone (O_3). 27 sites exceeded the CWS standards in the year 2002. Most of the sites (25) are in Quebec and Ontario. In British Colombia, the 98th percentile was in the range of 9-32$\mu g/m^3$ for the year 2002. Meanwhile, in Alberta the 98th percentile value was in the range of 8.0-25$\mu g/m^3$. However, Ontario and Quebec had a 98th percentile range of 23-41$\mu g/m^3$ and 26-38$\mu g/m^3$ respectively.

2.8 Conclusions

Particulate matter 2.5 ($PM_{2.5}$) consists of cocktail of chemical compounds. $PM_{2.5}$ is formed through wide suite of sources and reactions. Epidemiological studies have clearly shown that $PM_{2.5}$ causes heart and respiratory problems in humans. Also, due to its smaller size and its hygroscopic nature, $PM_{2.5}$ also causes visibility impairment in roadways. In Canada, achieving the Canada Wide Standards for $PM_{2.5}$ by 2010 has put the provinces to go for emission monitoring techniques.

2.9 References

Alberta Research Council (ARC), 1998. Overview on Particulate Matter. Report prepared for Alberta Environment. pp 70.

Bates, D. V., Baker Anderson, M., and Sizto, R., 1990. Asthma attack periodicity: a study of hospital emergency visits in Vancouver. *Environ. Res.*, **51**, 51-70.

Brunekreef, B., Dockery, D. W., Krzyzanowski, M., 1995. *Environ. Health Perspect* , **103**, 3.

Canadian Council of Ministers of the Environment (CCME) (2000). "*Canada-wide Standards for Particulate Matter and Ozone*", Guidance Document on Achievement Determination, Joint Action Implementation Coordinating Committee (JAICC).

CEPA/FPAC WGAQOG—Canadian Environmental Protection Act Federal-Provincial Advisory Committee Working Group on Air Quality Objectives and Guidelines. 1998. National Ambient Air Quality Objectives for Particulate Matter Science Assessment Document. Health Canada and Environment Canada, Ottawa and Toronto, ON.

Cheng, L., Sandhu, H. S., Angle, R. P., and Myrick, R. H. 1998. Particulate matter in Calgary and Edmonton. Emerging Air Issues for the 21st Century: The Need for Multidisciplinary Management: Proceedings of an International Specialty Conference. Air & Waste Management Association, Pittsburgh, PA.

Creason, J., Neas, L., Walsh, D., Williams, R., Sheldon, L., Liao, D.P., and Shy, C., 2001. Particulate matter and heart rate variability among elderly retirees: the Baltimore 1998 PM study. *J of Exposure analysis and Environmental Epidemiology*, **11** (2), 116-122.

Dockery, D. W., Pope, C. A., 3rd, Xu, X., Spengler, J. D., Ware, J. H., Fay, M. E., Ferris, B. G., Jr., and Speizer, F. E., 1993. An association between air pollution and mortality in six U.S. cities. *New England J of Medicine*, **329**(24): 1753-1759.

Dockery, D. W., and Pope, C. A., 1994. Acute Respiratory effects of Particulate Air Pollution. *Annual Rev. of Public Health*, **15**, 107-132.

Environment Canada 1995. Criteria Air Contaminant Inventory. Summary Report (Available online at www.ec.gc.ca/pdb/ape/cape_home_e.cfm)

Environment Canada, 1999. Canadian Environmental Protection Act, Priority Substances list Assessment Report. Respirable Particulate Matter Less Than or Equal to 10 Microns.

Environment Canada, 2000. Criteria Air Contaminants Emission Summaries for year 2000. (Available online at http://www.ec.gc.ca/pdb/ape/cac_ESummaries_e.cfm)

Environment Canada, 2004. *2002 National Summary of Ambient $PM_{2.5}$ and Ozone.* Report prepared for Joint action Implementation coordinating committee (Available online at http://www.cen-rce.org/eng/caucuses/energy/index.html)

Fang, M., Zheng, M., Wang, F., Chim, K.S., and Kot, S.C., 1999. The Long-range Transport of Aerosols from Northern China to Hong Kong—a multi technique study. *Atmos.Environ.*, 33, 1803–1817.

Hien, P. D., Bac, V. T., Tham, H. C., Nhan, D. D., and Vinh, L. D., 2002. Influence of Meteorological Conditions on $PM_{2.5}$ and $PM_{2.5-10}$ Concentrations during the Monsoon Season in Hanoi, Vietnam. *Atmos. Environ.*, **36**, 3473–3484.

Laakso, L., Hussein, T., Aarnio, P., Komppula, M., Hiltunen, V., Viisanen, Y. and Kulmala, M., 2003. Diurnal and Annual characteristics of Particle Mass and Number Concentrations in Urban, Rural and Arctic environments in Finland. *Atmos. Environ.*, **37**, 2629–2641.

Lebowitz, M.D., 1996. Epidemiological Studies of the Respiratory Effects of Air Pollution. *European Respiratory Journal*, **9**(5), 1029-1054.

Lippmann, M., and Thurston, G.D., 1996. Sulfate Concentrations as an Indicator of Ambient Particulate Matter Air Pollution for Health Risk Evaluations. *Journal of exposure Analysis and Environmental Epidemiology*, **6** (2), 123-146.

MacDonald, M., and Shepherd, 2004. M. Characterization of Visibility Impacts Related to Fine Particulate Matter in Canada. *J. Air and Waste Manage. Assoc.*, **54**(9), 1061-1068.

McKendry, I. G., Hacker, J. P., Stull, R., Sakiyama, S., Mignacca, D., and Reid, K., 2001. Long range transport of Asian Dust to the Lower Fraser Valley, British Columbia, Canada. *J. of Geophys. Res.-atm.*, **106**, 18361-18370.

Meng, Z., Dabdub, D., and Seinfeld, J.H., 1997. Chemical Coupling between Particulate Matter and Ozone. Science, **277**, 116-119.

Pandis, S. N., Wexler, A.S., and Seinfeld, J.H., 1995. Dynamics of Tropospheric Aerosols, *J. Phys. Chem.*, **99**, 9646.

Pathak, R.K., Yao, X., Lau, A.K.H. and Chan, C.K., 2003. Acidity and concentrations of ionic species of $PM_{2.5}$ in Hong Kong. *Atmos. Environ.*, **37**, 1113–1124.

Peters, A., Perz, S., Doring, A., Stieber, J., Koenig, W., Wichmann, H. E., 1999. Increases in heart rate during an air pollution episode. *American J of Epidemiology*, **150** (10), 1094-1098.

Pope, C. A., Verrier, R. L., Lovett, E. G., Larson, A. C., Raizenne, M. E., Kanner, R. E, Schwartz, J., Villegas, M., Gold, D. R., and Dockery, D. W., 1999a. Heart rate variability associated with particulate air pollution. *American Heart Journal*, **138** (5), 890-899.

Pope, C. A., Dockery, D. W., Kanner, R. E, Villegas, M., Schwartz, J., 1999b. Oxygen saturation, pulse rate, and particulate air pollution - A daily time-series panel study. *American J of Respiratory and Critical Care Medicine*, **159** (2), 365-372.

Sandhu, H.S. 1998. Ambient Particulate Matter in Alberta. Alberta Environmental Protection, Edmonton.

Schwartz, J., Dockery, D. W., and Neas, L. M., 1996. Is daily mortality associated specifically with fine particles? *J. Air Waste. Manage. Assoc.*, **46** (10), 927-939.

Seigneur, C. Air Pollution: Current challenges and Future Opportunities. (Available online at www.interscience.wiley.com).

Seinfeld, 1980. Lectures in Atmospheric Chemistry. American Institute of chemical Engineers. New York.

Seinfeld, J. H., and Pandis, S. N. 1998. *AtmosphericChemistry and Physics.* New York: Wiley. 1326 pp.

Seinfeld, J. H., and Pankow, J. F., 2003. Organic Atmospheric Particulate Material. *Annu. Rev. Phys. Chem.*, **54**, 121-140.

Seinfeld, J.H. Air Pollution: A Half-Century of Progress. (Available online at www.interscience.wiley.com).

Song, C. H., and Carmichael, G. R., 1999. The aging process of naturally emitted aerosols (sea-salt and mineral aerosol) during long range transport. *Atmos. Environ.*, **33** (14), 2203-2218.

Viotti, P., Luiti, G., and Di Genova, P. 2002. Atmospheric urban pollution: applications of an artificial neural network (ANN) to the city of Perugia. *Ecological Modelling*, **148**: 27-46.

Xu, X. P, Gao, Dockery, D.W., and Chen, Y., 1994. *Arch. Environ. Health* **49**, 216-219.

3.0 RECEPTOR MODELING TECHNIQUES AND ANN

3.1 Introduction

Stricter emission standards mean more pollution control. Receptor modeling techniques are normally used as a pollution control strategy. Receptor modeling techniques includes a wide range of multivariate analysis methods that use ambient air data to predict, forecast and identify source origins of pollutants (Watson and Chow, 2004). There are several receptor modeling techniques such as enrichment factor, chemical mass balance method, Eigen vector analysis, neural network, multiple linear regression, edge detection, cluster analysis and Fourier transform time series. Neural networks were used in this research because of their non-linear relationship characteristics and also because of their previous success in other fields such as, water, wastewater and weather forecasting. In this chapter a detailed description about different receptor modeling techniques, their advantages, disadvantages, and also a brief description about ANN and its success in pollution control has also been discussed.

3.2 Receptor modelling techniques

Receptor models have been used to quantify the source contribution of any pollutant in question. Figure 3-1 shows a source-receptor relationship. Monitors are receptors in addition to human, animals and materials. Readings taken from these monitors are used in the various receptor modeling techniques to identify the source emission.

Some of the receptor modeling techniques are as follows:

- Enrichment factor: Elemental enrichment factors (EFs) are widely used in environmental sciences to identify the sources of elements in air, atmospheric dust and precipitation (Reimann and Caritat, 2000). Moreover, they are also used in the air industry to distinguish between natural and anthropogenic sources of

metals (Berg et al., 1994). EFs is widely used receptor modeling technique to formulate pollution control strategies (Watson et al., 2002)

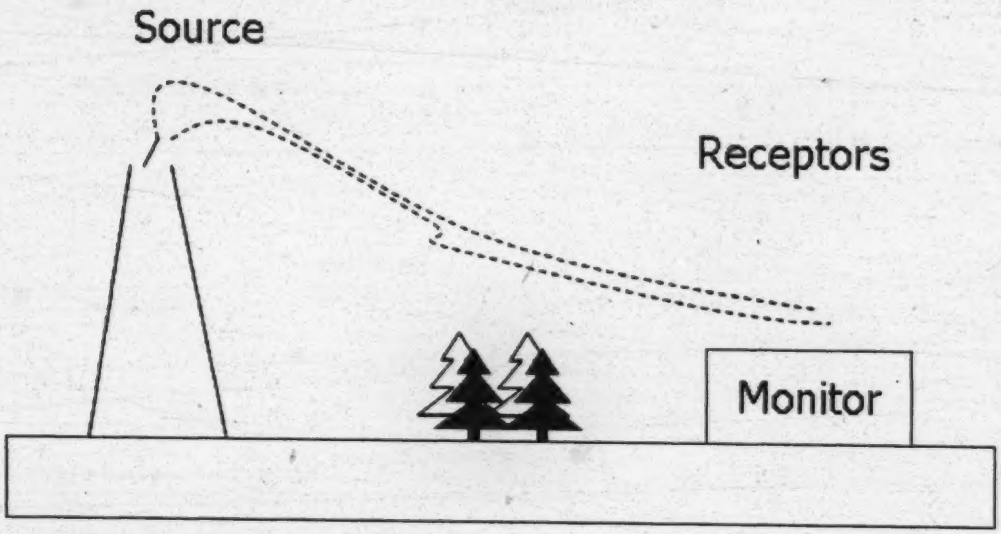


Figure 3- 1 Source-receptor relationship

- Multiple Linear Regression: This method attempts to determine the underlying relationship between set of inputs and a target output. Linear regression has been used in air quality applications (shi and Harrison, 1997; Ziomass et al., 1995). One of the disadvantages associated with linear regression is that, it cannot be used for non-linear systems (Gardner and Dorling, 1998).

- Multi-variate analysis: Multi-variate analysis is a popular statistical technique used for source identification of pollutants. Some of the common multi-variate techniques are principal component analysis (PCA), factor analysis and cluster analysis. The advantage of multi-variate analysis is that there is no priori knowledge about the study area is required. PCA, FA and Cluster analysis are qualitative techniques, which apportion sources based on relative not absolute contributions (Dias and Edwards, 2003; Watson and Chow, 2004).

- Chemical Mass Balance (CMB): CMB model calculates the major source contribution towards aerosol formation. Detailed source profiles are required for

finding out the source contribution, which is not possible because of the uncertainties associated with sampling and analytical techniques for a given source (Dias and Edwards, 2003). Like EFs, CMB is widely used to distinguish between anthropogenic and natural emission.

- Neural Networks: Neural networks are a mathematical form of receptor models. Receptor models are used for source identification and to study current state of pollution. Neural networks are widely used in various field to predict and forecasting purposes because of its ability to learn any complex non-linear relationships. Attempts are being made to use neural networks to identify pollutant sources. Lack of understanding on the working of neural networks hindered their acceptance for various applications. However, studies are undergoing to understand the "black box" like concept of neural networks (Zhang et al., 2004). In this research, neural networks were used because of the complexity in the formation of particulate matter.

3.3 Artificial Neural Networks – A theoretical background

Artificial Neural Networks (ANNs) are a versatile modeling technique for complex non-linear models (Zhang, 2002). It is different from standard computing techniques and from knowledge based expert systems in that instead of being programmed, it learns from the data representative of the problem to be solved (Serodes and Rodriguez, 1997).

ANNs are classified into feed-forward and feed back-ward networks. A feed-forward model allows signals to move from input to output nodes only (Static system). A feed-back model which is a dynamic system, output is generated for each input pattern (Jain et al., 1996). A feed-forward network is further classified into single-layer perceptron, multi-layer perceptron (MLP) and radial basis function (RBF) nets. On the other hand, feed-back networks are classified into competitive network, Kohonen network, Hopfield network, and ART models.

MLP is used in this research because of its previous success in the field of water (Baxter et al., 2002), wastewater (El-Din and Smith, 2002) and air (Hasham et al., 2004). A typical MLP consists of one or more layers of neurons configured in regular and highly connected topologies (Figure 3-2).

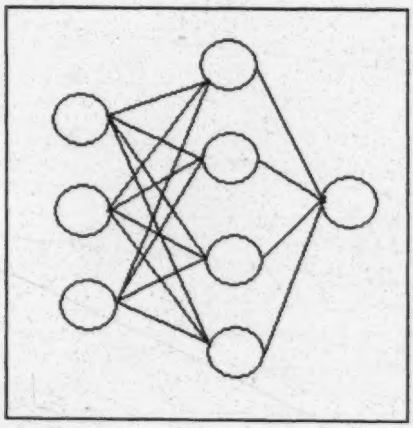


Figure 3- 2 A typical Neural Networks

Raw information is fed into the network via input nodes. Each neuron consists of an activation function and a set of related weights (Figure 3-3). Some of the common activation functions that is used in the hidden and output layer activation function were logistic, tangent, hyperbolic tangent, sine, Gaussian, Gaussian complement and symmetric logistic. There are no prior rules to arrive at the best activation function for the hidden and output layer. However, in this research a systematic approach was followed to arrive at the best activation function. Hasham et al., (2004) used a factorial design concept to arrive at the activation function of Gaussian-Logistic-Gaussian in 2-hidden layer and output layer respectively. In another study, Baxter et al., (2000) used logistic activation function both in hidden and output layer. Cannon and Lord (2000) used hyperbolic tangent and linear activation function in the hidden and output layer respectively.

Ability to learn through training makes ANNs attractive and exciting (Jain et al., 1996). There are three learning algorithms used in field of ANNs (Pal and Srimani, 1996):

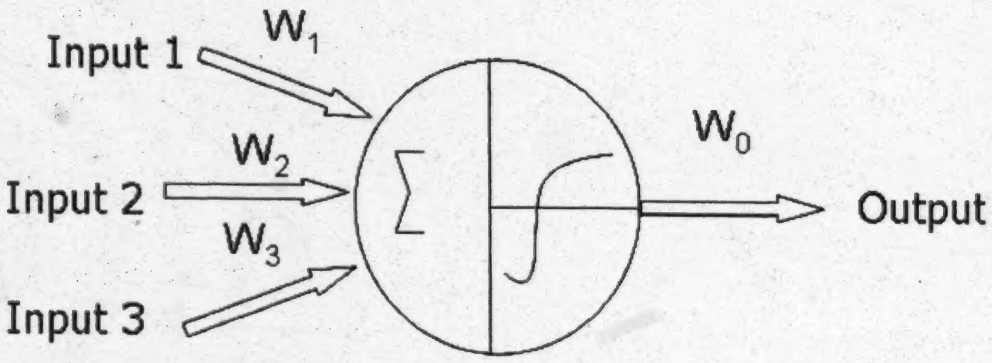


Figure 3-3 Internal working of a neuron

Supervised, unsupervised, and hybrid. Backpropagation algorithm (BP) which comes under the category of supervised learning algorithm and is widely used in the field of ANN modeling (Zhang and Stanley, 1997; Abdul-Wahab and Al-Alawi, 2002).In Hong Kong, Lu et al. (2004) used a combined principal component analysis (PCA), RBF model to forecast pollution trends. However, due to the lesser time taken in training the model, BP emerged as a viable alternative for our study. Cannon and Lord (2000) used a resilient back-propagation (RPROP) algorithm for training the model. RPROP adjusts its weight on its own while training.

In the BP algorithm, after an input has been applied to the neurons of input layers, it is propagated through each layer until an output is generated. This output is compared with the desired output, and an error signal is generated. This error is back-propagated to each neuron in the hidden layer. However, each layer receives a portion of the total error. Based on the error signal the connection weights are adjusted and the network training is continued until the error between the network output and desired output reaches a minimum. Care should be taken to avoid over-training of the model, which will make the network memorize the training patterns rather than mapping the input-output relationship.

Data patterns consisting of input and its related output are normally divided in to training, testing and production data sets. The training data set is used to train the model in understanding the relationship between input and output. The test data is used as a check,

so that during training of the network, the model is not caught in local minima (Gardner and Dorling, 1998). The production data set which is an independent data set is used to check the trained network's performance. There are several ways by which networks performance can be ascertained. Some of the statistical parameters which are used as a performance indicator (Su, 2004) are coefficient of multiple determinations (R^2), Pearson's product-moment correlation coefficient (r), Mean squared error (MSE), Mean absolute error (MAE), minimum absolute error, maximum absolute error, fractional bias (FB), Wilmott index of agreement (d_1 and d_2). In this research the coefficient of multiple determination (R^2) was used as a performance indicator.

3.4 ANNs and air pollution

Very few ANN studies for $PM_{2.5}$ as a class exist. Perez et al. (2000) modelled $PM_{2.5}$ concentrations in downtown Santiago, Chile. They compared their three layer multilayer perceptron model to a simple linear perceptron and a persistence model, finding that the multilayer perceptron outperformed the other two models. They used $PM_{2.5}$ concentrations in a 24-hour time window, relative humidity, and temperature as inputs to the model. The authors provided few statistical parameters to show their model performance. However, in 2001, Perez and Reyes repeated the study to forecast $PM_{2.5}$. This study found that when noise filtered data was input to the model, errors tended to accumulate for forecasts greater than one hour in advance.

McKendry (2002) compared ANN models to multiple linear regression models for predicting daily maximum hourly $PM_{2.5}$ concentration. It was found that of the variables he tested, the variables most important for $PM_{2.5}$ prediction were precipitation, wind speed, temperature, and season. Unlike the Perez models, McKendry used forecast meteorological variables in the models, increasing the potential for errors to occur.

In another study, Ordieres et al. (2005) compared the strengths and weakness of using MLP with RBF and Square multilayer perceptron (SMLP). They found that RBF had equally good performance like MLP, with lesser training time and greater network

stability. Moreover, they also found out that SMLP prediction did not have any particular advantage compared to that of MLP. They recommended RBF for replacing the traditional MLP.

He (2004) investigated the use of ANN in the Fort McKay area, Alberta, Canada. He used a Kohonen network to classify data according to meteorological data. He also used $PM_{2.5}$ concentration from nearby monitoring station as an input in addition to the meteorological and pollutant parameters from the Fort McKay station. The models developed had reasonable prediction capability compared to other models in the literature. The reason for models mediocre performance was speculated to be limited historical data. However, the author showed that pollutant parameters from other monitoring sites can be used as a predictor for the monitoring site in question.

Although there is a lack of $PM_{2.5}$ ANN models in the literature, there are many ANN models for PM_{10} (Chaloulakou et al., 2003), O_3 (Abdul-Wahab and Al-Alawi, 2002; Cannon and Lord, 2000; Gardner and Dorling, 2001; Guardani et al., 1999; Narsimhan et al., 2000, Spellman, 1999; Wotawa and Wotawa, 2001), SO_2 (Fernandez et al., 2003; Chelani et al., 2002), NO_2 (Gardner and Dorling, 1999; Hasham et al., 2004) and CO (Perez et al., 2004).

The lack of systematic approach in developing the models, non-availability of ample historical data, and exclusion of important input variables have hindered the prediction ability of most of the ANN models developed. In previous models developed, arbitrary (Gardner and Dorling, 2000; Cobourn et al., 2000), or thumb-rule approaches tend to be used in arriving at the ANN network architecture. Exclusion of input variables such as mixing height (Ruiz-Suarez et al., 1995; McKendry, 2002), precipitation data (Yaming, 2004; Perez et al., 2001) may be the reason for previously developed model's mediocre performance.

3.5 Conclusions

Neural networks were found to be a viable receptor modeling technique for Particulate matter, because of their ability to models complex non-linear situation. Some of the pitfalls associated in using ANNs are overcome in this research by following a systematic approach in model building.

3.6 References

Abdul-Wahab, S.A., and Al-Alawi, S.M., 2002: Assessment and prediction of tropospheric ozone concentration levels using Artificial Neural Network. *Environ. Modeling and Software*, **17,** 219-228.

Baxter, C.W., Stanley, S.J., Zhang, Q., and Smith, D.W. *Developing Artificial Neural Network Process Models: A Guide for Drinking Water Utilities*. in *6th Environmental Engineering Speciality Conference of the CSCE & 2nd Spring conference of the Geoenvironmental Division of the Canadian Geotechnical Society*. 2000. London, Ontario.

Berg, T., Royset, O., and Steinnes, E., 1994. Trace Elements in Atmospheric precipitation at Norwegian Background station (1989-1990) Measured by ICP-MS. *Atmos. Environ.*, **28**, 3519-3536.

Cannon, A.J., and Lord, E.R., 2000: Forecasting summertime surface level ozone concentrations in the Lower Fraser Valley of British Columbia: an ensemble neural network approach. *J. Air Waste Manage. Assoc.*, **50,** 322-229.

Chaloulakou, A., Grivas, G., Spyrellis, N., 2003. Neural Network and Multiple Regression Models for PM_{10} Prediction in Athens: A Comparative Assessment. *J. Air & Waste Manage. Assoc.*, **53**, 1183-1190.

Chelani, A.B., Rajghate, D.G., and Hasan, M.Z., 2002. Prediction of Ambient PM_{10} and Toxic Metals using Artificial Neural Networks. *J. Air & Waste Manage. Assoc.*, **52**, 805-810.

Dias, G.M., and Edwards, G.C., 2003. Differentiating Natural and Anthropogenic Sources of Metals to the Environment. *Human And Eco. Risk Ass.* **9**(4), 669-721.

Fernandez de Castro, B.M., Sanchez, J.M.P., Manteiga, W.G., Bande, M.F., Bermudez Cela, J.R., and Fernandez, J.J.H, 2003. Prediction of SO_2 Levels using Neural Networks. *J. Air & Waste Manage. Assoc.*, **53,** 532-539.

Gardner, M.W., and Dorling, S.R., 1999. Modeling and Prediction of Hourly NO_x and NO_2 Concentrations in Urban Air in London. *Atmos. Environ.*, **33,** 709-719.

Gardner, M.W., and Dorling, S.R., 1998: Artificial Neural Networks (The Multilayer Perceptron) - A review of applications in the atmospheric sciences. *Atmos. Environ.*, **32,** 2627-2636.

Guardani, R., Nascimento, C.A.O., Guardani, M.L.G., Martins, M.H.R.B., and Romano, J., 1999: Study of atmospheric ozone formation by means of a neural network-based model. *J. Air Waste Manage. Assoc.*, **49,** 316-323.

Hasham, F.A., Kindzierski, W.B., and Stanley, S.J., 2004. Modeling of Hourly NO_x Concentrations using Artificial Neural Networks. *J. Environ. Eng. Sci.*, **3,** 111-119.

He, Y. ANN Modeling of Ambient $PM_{2.5}$ in Fort McKay, Alberta. M.Sc Thesis, Department of Civil and Environmental Engineering, University of Alberta, Edmonton. 140pp.

Jain, A.K., and Mao, J., 1996: Artificial Neural Network: A Tutorial. *Computer*, **29,** 31-44.

Lu,W.Z., Wang, W.J., Wang, X.K., Yan, S.H. and Lam, J.C., 2004. Potential assessment of a neural network model with PCA/RBF approach for forecasting pollutant trends in Mong Kok urban air, Hong Kong. *Environ. Res.* , **96**(1), 79-87.

McKendry, I.G., 2002. Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM_{10} and $PM_{2.5}$) Forecasting. *J. Air & Waste Manage. Assoc.*, **52**, 1096-1101.

Narasimhan, R., Keller, J., Subramaniam, G., Raasch, E., Croley, B., Duncan, K., and Potter, W.T., 2000: Ozone modelling using neural networks. *J. Appl. Meteorol.*, **39,** 292-296.

Ordieres, J.B., Vergara, E.P., Capuz, R.S., and Salazar, R.E., 200. Neural Network Prediction Model for Fine Particulate Matter ($PM_{2.5}$) on the US-Mexico border in El-Paso (Texas) and Ciudad Judrez (Chihuahua). *Environmental Modelling and Software*, **20**(5), 547-559.

Pal, S.K., and Srimani, P.K., 1996. Neurocomputing - Motivation, models, and hybridization. *Computer,* **29**(3), 24-28.

Perez, P., Palacios, R., and Castillo, A., 2004. Carbon Monoxide Concentration Forecasting in Santiago, *Chile. J. Air & Waste Manag. Assoc.*, **54,** 908-913.

Perez, P., Reyes, J., 2001. Prediction of Particulate Air Pollution using Neural Techniques. *Neural Computing and Applications*, **10,** 165-171.

Perez, P., Trier, A., Reyes, J., 2000. Prediction of $PM_{2.5}$ Concentrations several hours in Advance using Neural Networks in Santiago, Chile. *Atmos. Environ.*, **34,** 1189-1196.

Rodriguez, M.J., Serodes, J.B., and Cote, P.A., 1997. Advanced chlorination Control in Drinking Water Systems Using Artificial Neural Networks. *Water Supply*, **15** (2), 159-168.

Reimann, C., and DE Caritat., 2000. Intrinsic Flaws of Element Enrichment Factors (EFs) in Environmental Geochemistry. *Environ. Sci. Technol.*, **34**, 5084-5091.

Spellman, G., 1999: An application of artificial neural networks to the prediction of surface ozone concentrations in the United Kingdom. *Appl. Geography*, **19,** 123-136.

Shi, J. P. and Harrison, R. M.: 1997, Regression modeling of hourly and concentrations in urban air in London, *Atmos. Environ.* **31**(24), 4081–4094.

Su, D. Application of Artificial Neural Network for Modelling Ground Level Ozone Concentrations in Calgary and Edmonton, Alberta. Department of Civil and Environmental Engineering, University of Alberta, Edmonton. 209pp.

Watson, J.G., and Chow, J.C., 2004. Receptor Models for Air Quality Management. Magazine for Environmental Managers, Air and Waste Management Association, Pittsburgh, PA, October 2004

Watson, J.G., Zhu, T., Chow, J.C., Engelbrecht, J., Fujita, E.M., and Wilson, W.E., 2002. Receptor Modeling Application Framework for Particle Source Apportionment. *Chemosphere*, 49, **9**, 1093-1136.

Wotawa, F., and Wotawa, G., 2001: Deriving qualitative rules from neural networks - a case study for ozone forecasting. *AI Communications*, **14,** 23-33.

Zhang, Q.J., Stanley, S.J., and Smith, D.W., 2004. Internal workings of Feed-Forward Neural Networks. *J. Environ. Eng. Sci.*, **3**, S1-S12.

Zhang, Q., 2002 Artificial Neural Network- Advanced Theories and Industrial Applications, Ph.D. Thesis, Department of Civil and Environmental Engineering, University of Alberta. Edmonton, pp 145.

Zhang, Q, and Stanley, S.J., 1997. Forecasting Raw-Water Quality Parameters for the North Saskatchewan River by Neural Network Modeling. *Wat. Res.*, **31** (9), 2340-2350.

4.0 EDMONTON EAST $PM_{2.5}$ MODELS

4.1 Introduction

Particles with aerodynamic diameters of 2.5 or less are called $PM_{2.5}$. $PM_{2.5}$ is formed due to anthropogenic (man-made) or biogenic (natural) sources. Epidemiological studies have linked $PM_{2.5}$ to respiratory (asthma) and cardiovascular problems (heart attack) (Burnett et al., 1998). Visibility, which is the prime public perception of air quality is also affected due to elevated $PM_{2.5}$ concentration (McDonald and Shepherd, 2004; Environment Canada, 1999). Currently the Canada Wide Standard (CWS) for $PM_{2.5}$ is set at $30\mu g/m^3$ (24-hour average) by the Canadian Government (Environment Canada, 1999).

Artificial Neural Networks (ANNs), an artificial intelligence (AI) technique, has the ability to generalize historical data patterns and make inferences about future trends. ANNs have been widely used in water treatment operations (Baxter et al., 2000), wastewater treatment (El-Din and Smith, 2002) and even prediction of change in weather patterns (Baawin et al., 2003) ANNs have also been used in the field of air pollution for the prediction of the levels of various pollutants such as ozone (O_3) (Abdul-Wahab and Al-Alawi, 2002), sulphur dioxide (SO_2) (Fernandez de Castro et al., 2003), oxides of nitrogen (NO_x) (Gardner and Dorling, 1999; Hasham et al., 2004) and particulate matter 10 (PM_{10}) (Chelani et al., 2002). However, prediction of $PM_{2.5}$ using ANNs has not been successful (McKendry, 2002; Perez and Reyes, 2001) perhaps due to the complex non-linear relationship between $PM_{2.5}$ and meteorology. Lack of a systematic approach in arriving at the proper architecture and seasonal variation observed in the concentration of $PM_{2.5}$ are some of the reasons for the poor generalization ability of the previously developed models.

A version of this chapter was submitted to the Air and waste Management conference which is going to be held in Minneapolis, Minnesota in June 2005. Also, the same paper is going to be submitted to the Journal of air and waste management.

$PM_{2.5}$ is a major threat to people and the environment, and its prediction in advance is important for environmental managers. The main objective of this research is to develop models for prediction (virtual monitor) and forecasting of $PM_{2.5}$ for the city of Edmonton, Canada using a systematic approach (El-Din and Smith, 2002). Analyzing the input variables that affect the concentration of $PM_{2.5}$ is the other objective of this research.

4.2 Artificial Neural Networks

Artificial Neural Networks (ANNs), an artificial intelligence (AI) technique have been used to understand complex non-linear mechanisms by repeated presentation of data patterns, several inputs, and its related output. One of the advantages of ANNs is that no priori knowledge of the study area is required. ANNs are classified into feed-forward networks and feed-back networks (Jain and Mao, 1996). Feed-forward networks are classified into single-layer perceptron (SLP), multi-layer perceptron (MLP) and radial basis function (RBF) networks. MLP has been widely used because of its success in various fields (Gardner and Dorling, 1998).

MLP consist of three or more layers of processing elements called neurons. Normally, there are three layers: input, hidden and output. Each layer has its own number of neurons. The Input layer processes the inputs and passes them on to the hidden layer. The hidden layer, of which there may be one or two, performs most of the processing of the three layers (Baxter et al., 2002) and passes on to the output layer. The output layer processes the information from the hidden layer and gives out predictions to the end user. Neurons, which are processing elements in an MLP is connected to each other through a set of weights. These weights are adjusted based on an error-minimization technique called back-propagation rule. Activation functions, which are present in the hidden and output layer produces output based on the summed weighted value passed to them. There are several activation functions: Logistic, tanh, Gaussian, Sine, tanh15, Gaussian complement and symmetric logistic.

Training of the model is done using a supervised learning methodology, presenting of historical data patterns, consisting of a set of inputs and its related output. Training is

continued until the error between the desired output and actual output is reduced to a minimum. The back-propagation rule adjusts the weights to minimize the error. Test set data are used as a check to assess the model's training so that the model does not memorize the interactions. The trained model is applied to the production data set, which is independent of the training data set, to ascertain the performance of the model. The model should be trained until the generalization performances reach a maximum. One way to avoid over training is by using the early stopping technique.

4.3 Data and methods

4.3.1 Study Area

Edmonton, the capital of the province of Alberta, Canada is located at 53.5° N latitude and 113.5° W longitude. Edmonton, a northern city in the province of Alberta, Canada, is characterized by short summers and long winters. Ground-based inversions are a common phenomenon in this city (Sakiyama et al., 1994). Edmonton has gas-fired power plants, petroleum refineries, cement kilns, coal-fired power plants and asphalt roofing manufacturing plants (Cheng et al., 1998). There are also cases of forest fires, which enhance the level of air pollutants in the atmosphere in Edmonton (Cheng et al., 1998a).

Air pollutants are monitored on a continuous basis by three monitoring stations: East, Central and North. In this research, data from the East monitoring station were used for model development. Mixing height (twice daily) and opacity data were obtained from the Stony Plain Station, 20km west from Edmonton.

4.3.2 Model Inputs

The most important criteria in developing the model were choosing the input variables which might affect the output. The following are the input variables that were included in developing the models. Average hourly data for pollutants and atmospheric variables were used in the model development based on previous studies done in the field of air pollution.

Carbon monoxide (CO). In Edmonton, the transportation sector contributed 48% of the total mass of $PM_{2.5}$ (Cheng et al., 1998). CO is one of the pollutants emitted from the vehicle related emissions. It is included as an input for the prediction of $PM_{2.5}$, because of their linear correlation with $PM_{2.5}$ (Perez et al., 2004; Bogo et al., 2003). It is also used as an indirect measure of atmospheric stability and WSP (Abdul-Wahab and Al-Alawi, 2002).

Oxides of nitrogen (NO_X). NO_x is the sum of NO and Nitrogen dioxide (NO_2). Most Primary NO_x is emitted in the form of NO (Gardner and Dorling, 1999).Vehicle and combustion processes are the major sources of NO_x. In Alberta, forest fires emit significant amount of NO_x (Cheng et al., 1998a).

NO reacts with ozone and forms the secondary NO_x, NO_2. There is a strong correlation between $PM_{2.5}$ and NO_X (Harrison et al., 1997). Both NO and NO_2 were included as a separate input variable in developing the model.

Ozone (O_3). O_3 is a secondary pollutant, formed through a reaction between NO_x and volatile organic compounds (VOCs) in the presence of sunlight. O_3 is included as an input to simulate complex atmospheric chemistry.

Sulphur dioxide (SO_2). In Edmonton, sulphate mass contributed 11% of the total mass fraction of $PM_{2.5}$ (Cheng et al., 1998).

Total hydrocarbon (THC). Vehicles are the major sources of hydrocarbons. Carbonaceous compounds are a major constituent in the mass of $PM_{2.5}$ (Harrison et al., 1997).

Temperature (Temp). Temp acts as a major indicator of seasonal change. In previous study by McKendry (2002) on $PM_{2.5}$, temperature was found to be an important variable. There is a strong relation between high ozone episodes and high temperature (Sandhu, 1998).

Mixing height (MH). Inclusion of MH simplifies complex meteorological processes, which acts as a ceiling in trapping air pollutants.

Wind speed (WSP). There is a strong inverse correlation between WSP and $PM_{2.5}$ concentration (Harrison et al., 1997 and Chaloulakou et al., 2003). WSP plays an important role in the distribution of pollutants over an area. Many of the pollution episodes occurred only on days during which convective conditions when wind speed levels were low.

Wind direction (WDR). WDR indicates the direction from which the wind blows, and is measured in degrees from north. Moreover, many studies in the field of air pollution modeling used WDR as an input parameter in modeling (Chaloulakou et al., 2003a).

Wind direction deviation (DEV). DEV may be used as a measure of atmospheric stability (Hasham et al., 2004). DEV has been used as a description of atmospheric turbulence in dispersion models (Weber, 1997).

Opacity (OPA). OPA may be used as a surrogate for solar radiation received. Solar radiation plays an important role in photochemical enhanced reactions. Particulate concentration is higher during sunny days due to enhanced photochemical activity (Chaloulakou et al., 2003).

Relative humidity (RH). RH may serve as a surrogate parameter for precipitation as surface wetness controls the concentration of $PM_{2.5}$. Studies (McKendry, 2002; Perez and Reyes, 2001) done on $PM_{2.5}$ have used RH as an input parameter.

Month of the year. Month of the year was included as a separate input to account for seasonality in the concentration of $PM_{2.5}$ (Su et al., 2004). Indexing of variables was used in this research for developing the models. For example, if the data pattern was for the month of January, then an index of 1 was assigned to it and 0 was assigned to all other months.

Day of the week. Historical study on the data trends in Edmonton and elsewhere in the world has shown that, concentration of $PM_{2.5}$ is higher on weekdays compared to weekends (Su et al., 2004, 2004a; McCullum et al., 2004, 2004a; Melas et al., 2002). To account for this variation, an indexed variable for day of the week is used.

Hour of the day. Diurnal variation in concentration of $PM_{2.5}$ was observed in Edmonton (Su et al., 2004; McCullum et al., 2004). Indexing of hour of the day was used to account for this variation.

4.3.3 Model Development

Model development was done according to a systematic approach followed in wastewater treatment modelling (El-Din and Smith, 2002). An optimum network structure was obtained by following a systematic approach of model development (Su et al., 2004b). Hourly pollutant and meteorological data from the Edmonton East station were used in the model development. MH data was calculated in two ways: Interpolation between twice-daily balloon sounding data measured at Stony Plain station and the use of three-hour centered averaged WSP data (Benkley and Schulman, 1979). Lower values of MH obtained by two methods were used in developing the models. To differentiate between 0° and 360°, the wind direction 0° was changed to 360°. In developing the models, historical data from the year September 2000 to August 2003 were used. Ward systems Group, Inc.'s Neuroshell 2 (Release 4.0) software and its associated batch processor feature were used exclusively to develop the models.

4.3.4 Systematic Approach

1. Historical data relating to the modeled process are obtained and pre-processed to remove erroneous entries. In Alberta, publicly available ambient air monitoring data has generally been purged of erroneous entries prior to being made public. However, some hours have missing data.

2. Statistics for the data are calculated to characterize each parameter. The data are plotted to visually identify any temporal trends or anomalies.

3. The data are subdivided into training, test, and production data subsets in a 3:1:1 ratio. Temporal inputs such as month, day, and hour are converted to numerical format (i.e. for month, January =1, February = 2, etc.).

4. The most appropriate activation functions for the hidden and output layers are determined through testing of various combinations. The activation function combinations are evaluated at three settings for the number of neurons in the hidden layer and the number of training epochs. These settings, called the low, middle, and high settings for simplicity, are selected with the expectation that the final network structure will fall within this range. The number of hidden layer neurons and training epochs at each setting are described in Table 4-1. The activation functions evaluated for the hidden layer were the logistic, hyperbolic tangent (tanh), hyperbolic tangent 1.5 (tanh1.5), Gaussian, Gaussian complement, sine, and symmetric logistic functions. The activation functions tested for the output layer included all those tested for the hidden layer and the linear function.

5. The performance of each activation function combination is evaluated to determine the best activation function combination averaged over all three settings. The combination selected must be valid at all three settings to ensure the final network is stable regardless of its structure.

6. The number of neurons in the hidden layer that best describes the process and the number of training epochs are determined. Each feature is varied through a range of values and combinations to arrive at the best performing model, evaluated using the model coefficient of multiple determination, R^2. Where two models have comparable performance, the model requiring the least runtime, contains the least number of hidden layer neurons, or requires the

fewest number of epochs to train, is preferred according to the principle of parsimony.

7. The most relevant input variables are identified using their relative input contributions determined by the ANN. The current scientific knowledge of the process is also taken into consideration. Comparing the model's performance when the parameter is included and excluded from the model inputs verifies the parameter's importance.

8. Inclusion of the previous hours' $PM_{2.5}$ concentrations in the model is evaluated to account for persistence effects.

9. An analysis of the residuals for the best network is completed to identify any biases. The model's stability is confirmed by swapping the data in the training, test, and production sets and re-training the network. In this process, the test and production set data are re-assigned to the training set, while two thirds of the training data are apportioned between the test and production sets. If the model is stable, the retrained network should perform at a level similar to the original network.

Table 4- 1 Number of hidden layer neurons and training epochs setting for evaluating hidden and output layer activation function.

Setting	No. of neurons in hidden layer	No. of training epochs
Low	5	500
Middle	20	4000
High	40	7500

4.4 Results and discussion

A systematic approach outlined in the previous section was followed in developing the models. The virtual monitor was developed first, and the complete set of historical data after error pre-processing was used to develop it. The results obtained were poor, with the

October, and November) and winter (December, January, and February), and models were developed for all the seasons. Table 4-2 shows the basic statistic for all the seasons.

4.4.1 Virtual Monitor

Season specific models were developed because of the poor predictive ability of the models developed with full year data. Activation functions were determined for the hidden and output layer for all the seasons based on the 4th step in the systematic approach. Table 4-3 shows the activation function which was found using step 4 of the systematic approach. Using this activation functions models were developed for all seasons.

Table 4- 2 Basic statistics for different season historical data.

(a) Fall season

Parameters	Mean	SD	Max	99th percentile
CO (ppm)	0.37	0.27	3.10	1.40
NO (ppm)	0.02	0.03	0.39	0.16
NO2 (ppm)	0.02	0.01	0.07	0.05
O_3 (ppm)	0.02	0.01	0.06	0.04
SO_2 (ppm)	0.00	0.00	0.04	0.01
THC (ppm)	2.42	0.85	24.20	5.47
MIX (m)	129.72	169.05	953.70	659.40
OPA (tenths)	4.49	3.96	10.00	10.00
RH (%)	69.98	19.44	100.00	100.00
TEMP (°C)	5.50	7.93	32.70	24.50
WDR (°)	208.58	84.76	360.00	354.00
DEV(°)	18.61	16.86	166.00	98.00
WSP(km/h)	9.51	4.95	30.30	23.57
$PM_{2.5}$ (µg/m^3)	8.04	7.01	72.40	33.64

(b) winter season

Mean	SD	Max	99th percentile
0.48	0.32	2.80	1.70
0.03	0.04	0.48	0.21
0.03	0.01	0.08	0.06
0.01	0.01	0.05	0.04
0.002	0.00	0.04	0.01
2.55	0.80	12.10	6.30
74.94	126.45	965.28	538.19
4.56	4.10	10.00	10.00
76.54	13.91	100.00	99.00
-7.27	8.15	14.00	7.20
200.78	83.58	360.00	357.00
15.96	15.15	163.00	93.00
8.87	4.69	36.90	21.90
8.59	8.39	123.30	38.00

(c) Spring season

(d) summer season

Parameters	Mean	SD	Max	99th percentile	Mean	SD	Max	99th percentile
CO (ppm)	0.33	0.18	1.8	1.00	0.25	0.13	1	0.70
NO (ppm)	0.01	0.02	0.20	0.10	0.01	0.01	0.14	0.06
NO2 (ppm)	0.02	0.01	0.08	0.06	0.01	0.01	0.06	0.04
O_3 (ppm)	0.03	0.02	0.07	0.06	0.03	0.02	0.10	0.07
SO_2 (ppm)	0.00	0.00	0.06	0.01	0.00	0.00	0.03	0.01
THC (ppm)	2.3	0.67	13.1	4.90	2.33	1.06	25	6.50
MIX (m)	261.28	203.46	2170.58	826.39	220.07	182.18	2261.5	716.18
OPA (tenths)	4.61	3.88	10.00	10.00	3.98	3.63	10.00	10.00
RH (%)	59.87	22.64	100.00	100.00	60.59	22.96	100.00	100.00
TEMP (°C)	2.53	11.04	30.6	25	18.33	6.15	38	33.34
WDR (°)	200.11	96.42	360.00	357	206.46	97.67	360.00	357.00
DEV(°)	17.80	14.16	168.00	82.69	23.46	19.69	166.00	116.84
WSP(km/h)	10.50	5.78	35.80	26.2	8.78	5.24	30.9	23.10
$PM_{2.5}$ ($\mu g/m^3$)	7.51	11.12	295.0	37.63	9.19	7.81	80.8	36.88

Table 4- 3 Summary of architecture followed in developing seasonal virtual monitor models.

Season	Hidden layer activation function	Output layer activation function
Spring	Sine	Sine
Summer	Tanh	Gaussian complement
Fall	Logistic	Logistic
Winter	Gaussian complement	Logistic

4.4.1.1 Spring model

The spring season is optimized for best architecture as outlined in the step 5 of the systematic approach. Accordingly, the best architecture was found to be 5 hidden layer

neurons and 500 training epochs (Figure-4-1). The R^2- value for the best performing model for the spring season was 0.72.

The model employed the Sine activation function in both the hidden and output layers, and trained in 3:47 minutes. The model performance shown in Figure 4-2 is typical of most $PM_{2.5}$ models, slightly under-predicting peak values. The spring virtual monitor model had 5 neurons in the hidden layer and was trained with 500 epochs. Residual check performed on the spring data did not show any obvious trends (Figure A-1 and A-2).

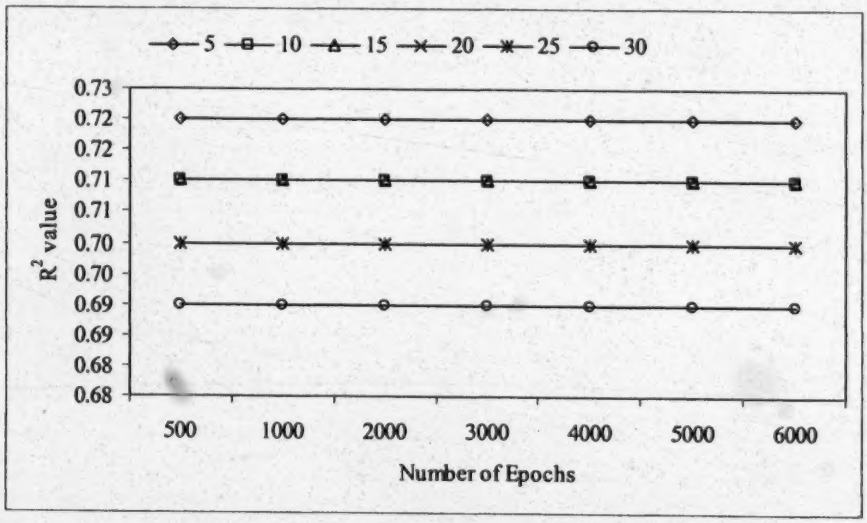


Figure 4- 1 Optimization of model architecture for spring season virtual monitor model.

4.4.1.2 Summer season model

Figure 4-3 shows the optimization of summer season model architecture. The best model architecture was found to be 500 training epochs and 20 hidden layer neurons with a R^2- value of 0.41.

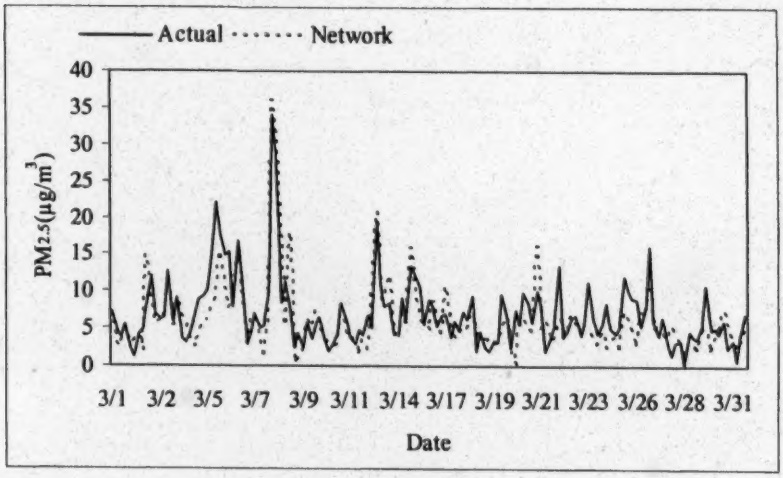


Figure 4- 2 Edmonton East spring virtual monitor model. $R^2 = 0.65$.

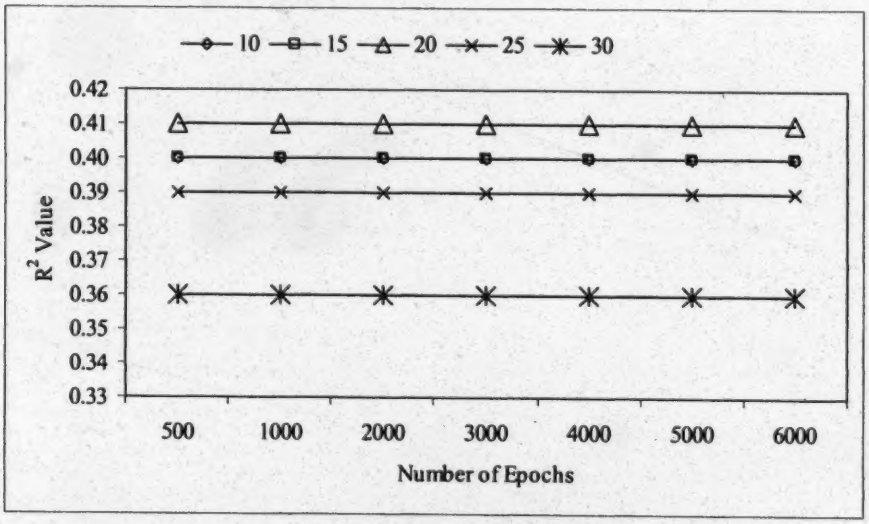


Figure 4- 3 Optimization of model architecture for summer season.

The summer model showed poor performance compared to the other seasons. Within the historical data, few events where the $PM_{2.5}$ concentration was greater than 50 $\mu g/m^3$ occurred. The peaks where the $PM_{2.5}$ concentrations were greater than 50 $\mu g/m^3$ may be due to forest fire effects.

The model developed with 20 hidden layer neurons, a hyperbolic tangent activation function in the hidden layer, a Gaussian complement activation function in the output

layer, and trained with 500 epochs resulted in a production set R^2 of 0.41. In an attempt to improve the model performance, rainfall data in the format of a rain index (0 = no rain, 1 = light rain, 2 = moderate rain, and 3 = heavy rain), were included as an input variable. There was no improvement to the model performance with the addition of rainfall data. An example of the model predictions compared to actual values, excluding rainfall data, is provided in Figure 4-4. In general, the model was able to follow the $PM_{2.5}$ concentration trend, but failed to capture the extent of peaks. Residual analysis performed on the summer season virtual monitor model did not show any obvious trends (Figure A-3 and A-4).

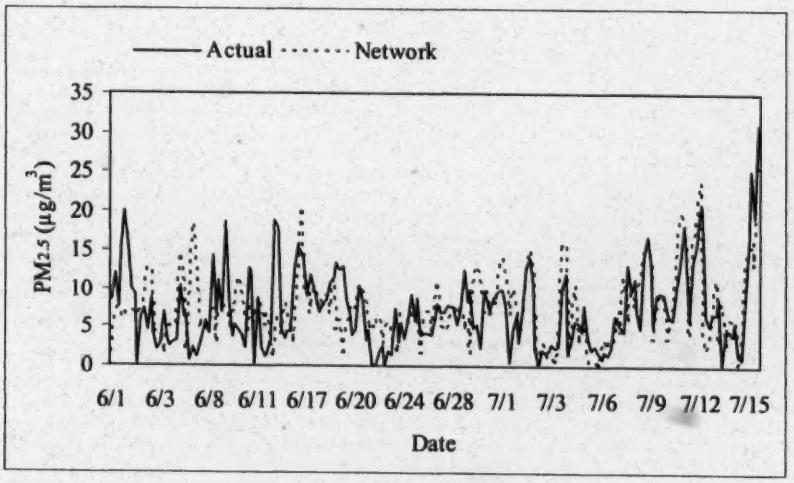


Figure 4- 4 Edmonton East summer virtual monitor model. $R^2 = 0.41$.

4.4.1.3 Fall season

The fall season virtual monitor model was optimized using logistic activation function in the hidden and output layer (Figure 4-5). The fall model was developed with 3737 data patterns in the training set, 1246 patterns in the test set, and 1246 patterns in the production data set. $PM_{2.5}$ ranged from a minimum of 0 to a maximum of 72.4 $\mu g/m^3$ for the historical data used in this study. The combination of the logistic activation function in both the hidden and output layers performed best, producing a production set R^2 of 0.66, when a network of 30 neurons in the hidden layer and 500 training epochs was used. The model performance on a sampling of the production data set is shown in Figure 4-6.

The ANN was able to predict the $PM_{2.5}$ trends well, neither consistently over- or under-predicting the peaks.

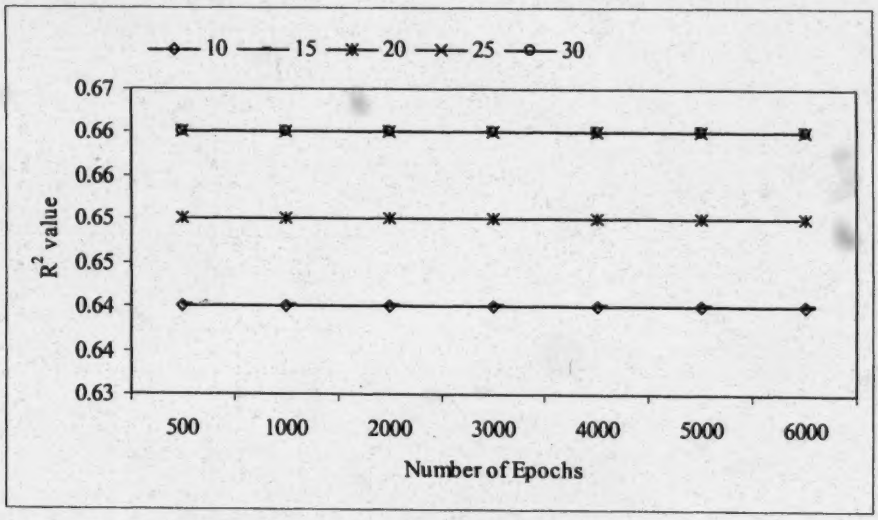


Figure 4- 5 Optimization of model architecture for fall season.

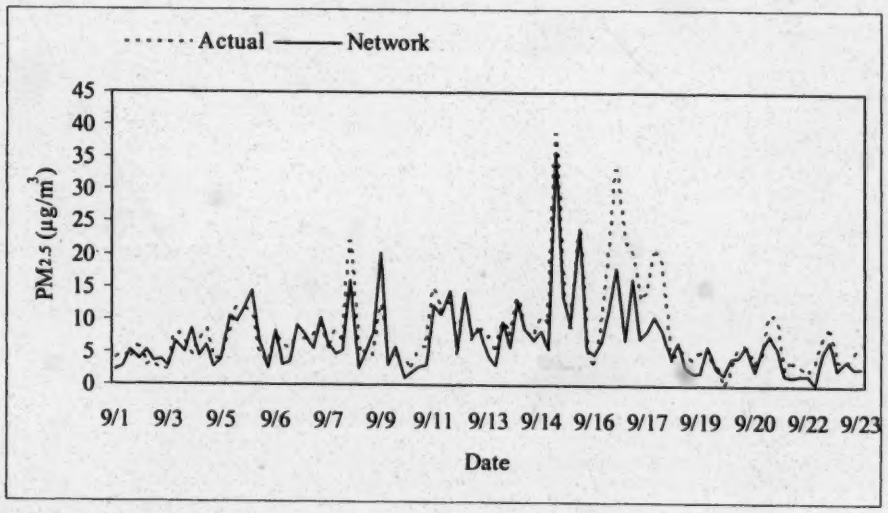


Figure 4- 6 Edmonton East fall virtual monitor model. $R^2 = 0.72$.

4.4.1.4 Winter season

The average $PM_{2.5}$ concentration in the winter was 8.5 $\mu g/m^3$, with a range from 0 to 123 $\mu g/m^3$. Figure 4-7 shows the optimization of network architecture. Accordingly, with a network built of 15 hidden layer neurons, a Gaussian complement activation function in

the hidden layer and a logistic activation function in the output layer, the model R^2 on the production set was 0.72. The model was trained in 500 epochs. The model predictions are shown in Figure 4-8 As with the other models, trends were predicted well. Table-4 summarizes the virtual model characteristics for all the seasons. The Residual analysis showed no discernible trends or biases in the prediction which is evident from the Figures A-7 and A-8.

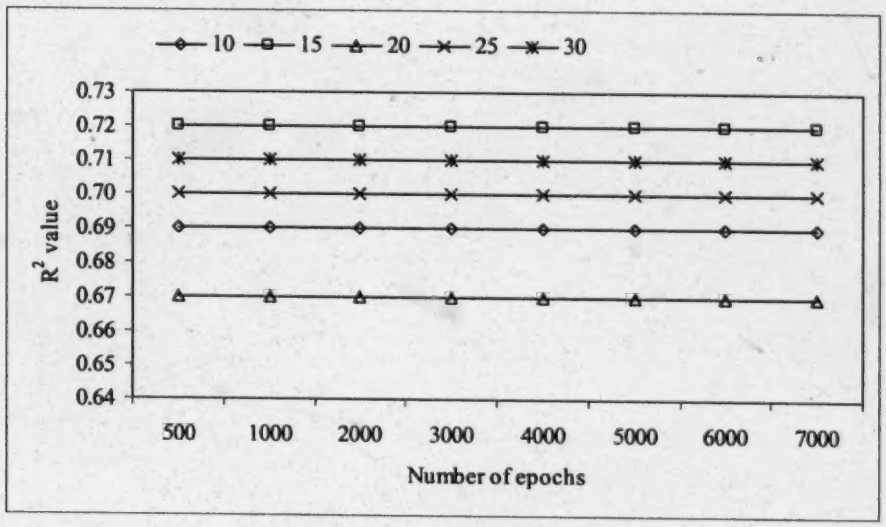


Figure 4- 7 Optimization of model architecture for winter season.

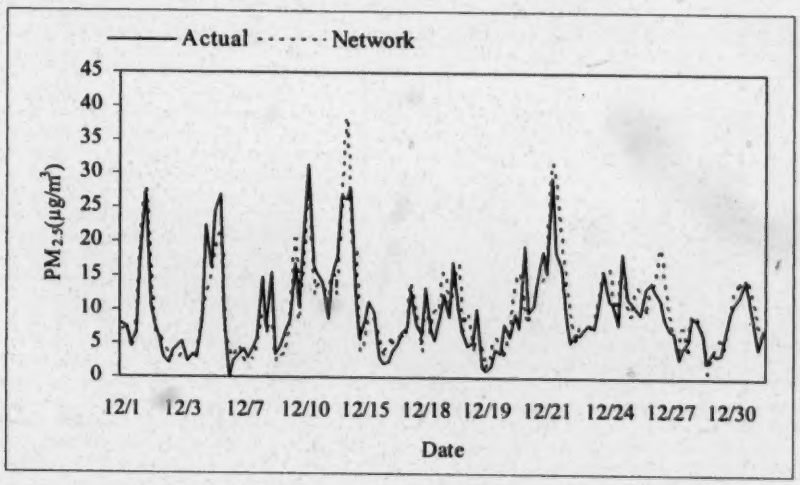


Figure 4- 8 Edmonton East winter virtual monitor model. $R^2 = 0.72$.

Table 4- 4 Summary of architecture followed in developing seasonal virtual monitor models.

Feature	Spring	Summer	Fall	Winter
No. neurons in hidden layer	5	20	30	15
Hidden layer activation function	Sine	tanh	logistic	Gaussian complement
Output layer activation function	Sine	Gaussian complement	logistic	Logistic
No. training epochs	500	500	500	500
Training runtime (min: sec)	1:04	2:43	5:04	5:25
No. training set patterns	3379	2555	3737	3782
No. test set patterns	1127	852	1246	1261
No. production set patterns	1126	852	1246	1260
Production set R^2	0.72	0.41	0.66	0.72

4.4.2 Important input variables

$PM_{2.5}$ is formed through a complex mechanism between pollutants and meteorology. Variables that affect the formation of $PM_{2.5}$ were found by removing the variables one at a time and keeping the pollutant parameters CO, NO, NO_2, SO_2, O_3, THC and indexed variables as constant based on the theoretical background for $PM_{2.5}$. Any drastic change in R^2 value is because the result of that variable's importance in prediction. Table 4-5 shows the change in R^2 value on removing a particular variable.

RH was found to be important for the winter and fall seasons due to the high mean RH prevalent in winter (77%) and fall (70%) seasons (Table 4-2) which acts as a sink in removing particles. MH was found to have minimum effect in the overall

Table 4- 5 Change in R^2 value with respect to removal of particular variable

	Spring	Summer	Fall	Winter
Initial R^2 value	0.72	0.41	0.66	0.71
OPA	0.47	0.45	0.60	0.70
RH	0.71	0.42	0.61	0.67
WDR	0.71	0.40	0.63	0.69
MH	0.69	0.40	0.64	0.69
WSP	0.65	0.37	0.65	0.69
TEMP	0.44	0.37	0.61	0.66
DEV	0.53	0.41	0.62	0.70

prediction for all the seasons. This may be due to the disadvantages associated with the calculation of hourly MH data. Opacity which acts as a surrogate for solar radiation received emerged as an important variable in summer, fall, and spring season. One reason for this may be due to the photochemical - associated particulate formation in these seasons. Temp was found to be important for all seasons.

WDR was found to be important for the fall and winter seasons, whereas WSP was found to be unimportant during the same period of time. The location of cattle farms on the southern side of Edmonton and the predominant WDR from the south may be the reason for the importance of WDR in fall and winter.

The importance of WSP in the summer season may be due to the turbulent conditions prevalent in the area, was evident from the mean WDR deviation value (23°), which comes under the category of unstable atmospheric condition (EPA, 1986; Mitchelle and Timbre, 1979). Transboundary transport of pollutants is associated with turbulent conditions. There was a case of transboundary intrusion of dust particles in the Lower Fraser valley in the neighboring Province of British Columbia (McKendry et al., 2001). However, standard deviation of wind direction was found to be unimportant for the summer season.

From Table 4-2, the inverse correlation between WSP (10.5km/hr) and $PM_{2.5}$ concentration (7.5µg/m^3) was observed to be higher for the spring season. According to Chaloulakou (2003a), emissions from local sources, dominate due to the strong inverse correlation. This may be the reason for WSP's importance in the spring season. Table 4-6 summarizes the key variables to the prediction of $PM_{2.5}$

4.4.3 Forecast models

In forecast models, the $PM_{2.5}$ concentrations were predicted 1-hr in advance. Compared to the virtual monitor models, this model uses present $PM_{2.5}$ concentration as an input. The hidden layer and output layer activation function that was used in developing the virtual monitor models was also used in the forecast models.

4.4.3.1 Spring season

The Edmonton spring forecast model architecture consists of 10 hidden layer neurons. The network required 500 epochs and 11:57 minutes to train. The resulting network performed at a R^2 of 0.65 for the production set of data. The network performance is shown in Figure 4-9. For large peaks, the network tended to under-predict $PM_{2.5}$ concentrations. Small and mid-level peaks were generally overshot. Figure A-9 in Appendix-B shows the plot of Network versus Predicted $PM_{2.5}$ concentration. From the Figure it can be ascertained that the points are distributed equally on both sides of the solid black line. Also, the residual plot (Figure A-10) shows that there are no obvious trends. Thus, spring forecast model appears to be a stable one.

4.4.3.2 Summer season

The summer forecast model performed better than the virtual monitor model, producing a R^2 of 0.58 for the production set. The ANN was able to predict the general $PM_{2.5}$ trend, but over-predicted some mid-level peaks and under-predicted high peaks. The model was trained in 500 epochs and 1:28 minutes. The performance of summer forecast model is shown in Figure 4-10. Residual analysis performed on the summer model, showed some sort of pattern (Figure A-11 and A-12).

Table 4- 6 Input variables important in the prediction of $PM_{2.5}$ for all the seasons

Input variable (No. of indexed input)	Season			
	Fall	Winter	Spring	Summer
CO	✓	✓	✓	✓
NO	✓	✓	✓	✓
NO_2	✓	✓	✓	✓
SO_2	✓	✓	✓	✓
THC	✓	✓	✓	✓
O_3	✓	✓	✓	✓
Month (12)	✓	✓	✓	✓
Day (7)	✓	✓	✓	✓
Hour (24)	✓	✓	✓	✓
Temp	✓	✓	✓	✓
RH	✓	✓	✓	
Opacity	✓		✓	✓
WSP			✓	✓
WDR	✓	✓		
DEV	✓		✓	
MH				

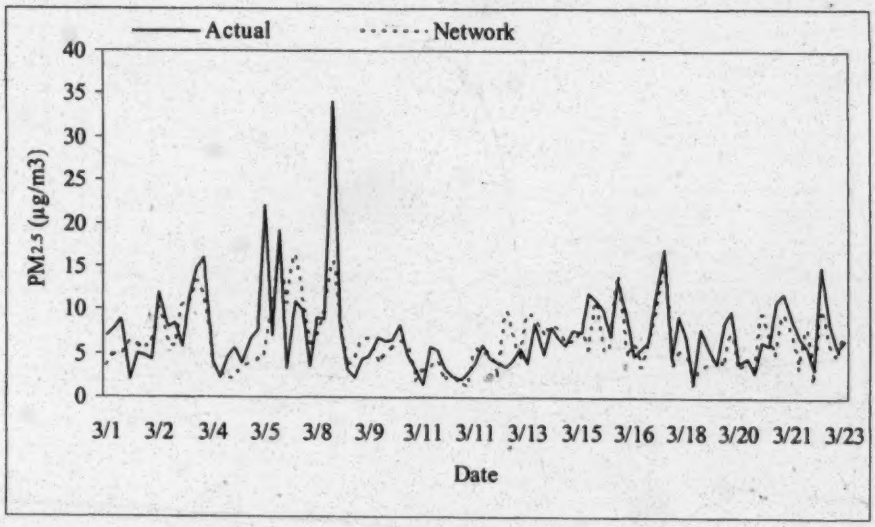


Figure 4- 9 Edmonton East spring forecast model. $R^2 = 0.72$.

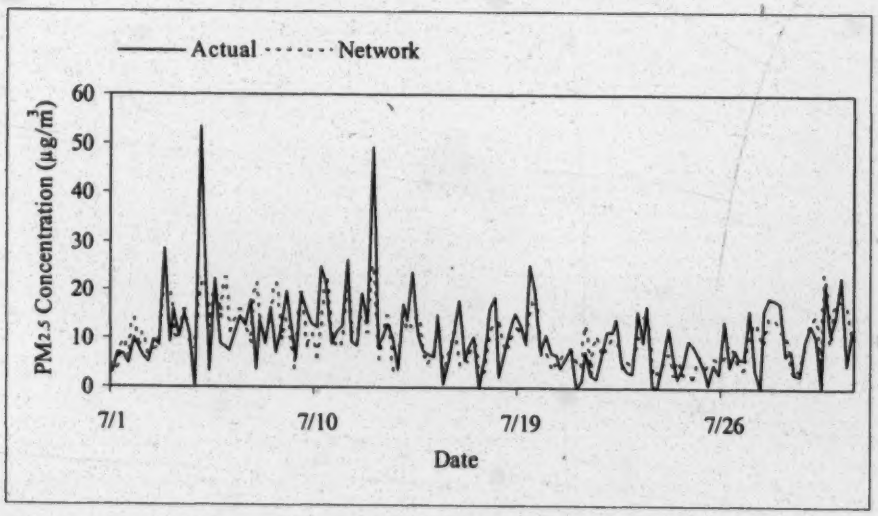


Figure 4- 10 Edmonton East summer forecast model. $R^2 = 0.58$

4.4.3.3 Fall season

The fall model performed at an equivalent level to the spring forecast, with a R^2 of 0.76 for the production data set. The fall forecast model contained 10 neurons in the hidden layer, and required 500 epochs to train. The total runtime needed for training the model was 3:19 minutes. The fall forecast model performance is shown in Figure 4-11. The figure shows that the fall model predicted the $PM_{2.5}$ concentration trends well, with no

large over- or under-predictions. Residual analysis showed an evenly scattered distribution of residuals (Figure A-13 and A-14).

4.4.3.4 Winter season

The ANN forecast model for the winter season was similar to that of the fall model, resulting in a R^2 of 0.78 for the production set of data. The winter forecast model had 25 neurons in the hidden layer. The model required 500 training epochs and a runtime of 4:30 minutes to train. The winter model performance is shown in Figure 4-12. The model had a combination of over- and under-predictions for mid-level and low-valued peaks.

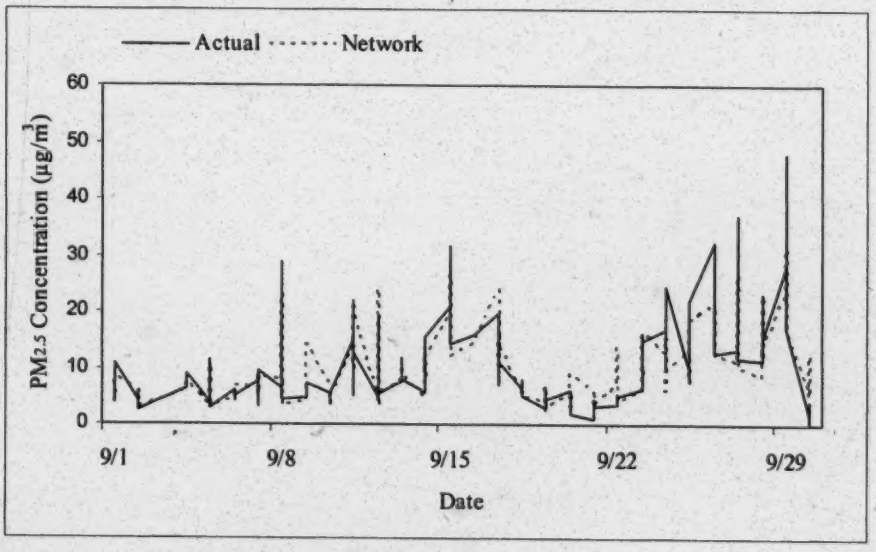


Figure 4- 11 Edmonton fall forecast model. $R^2 = 0.76$.

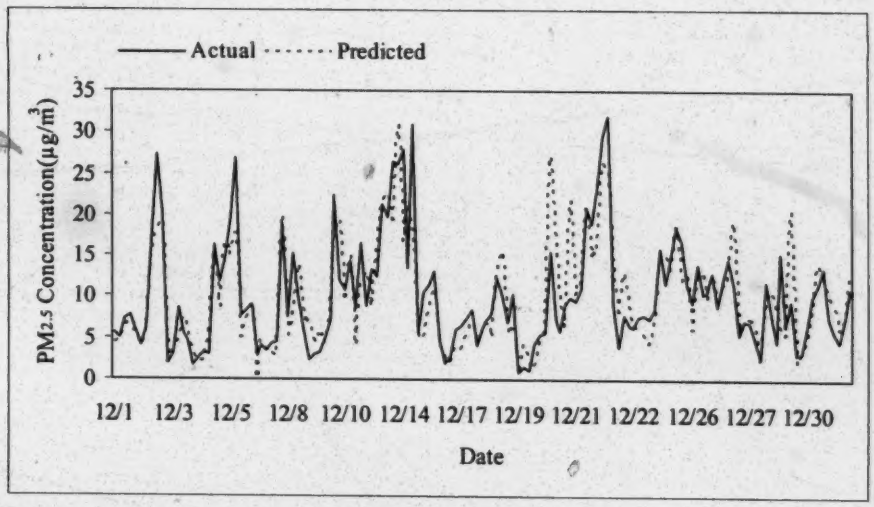


Figure 4- 12 Edmonton winter forecast model. $R^2 = 0.78$.

Overall, the forecast models performed better than the virtual monitor models for all the seasons except the spring. The winter season had the highest prediction capability compared to the other season forecast models with an R^2 value of 0.78. The fall season forecast model, with an R^2 value of 0.76, was able to predict the lows of $PM_{2.5}$ concentration but had difficulty in predicting the higher trends. Meanwhile, the summer season model, with an R^2 value of 0.58, was able to predict the low values of $PM_{2.5}$ concentration better than the higher values. However, the prediction ability of the spring season forecast model decreased from a R^2 value of 0.72 to 0.65, with the addition of present hour $PM_{2.5}$ concentration, which is contrary to the other seasonal models.

4.4.4 Check for persistence

The addition of persistence (lagged data) of data as an input and retraining the model were carried out as a part of the systematic approach of model building. In one study done in Athens (Chaloulakou et al., 2003), pollutant levels during the previous three days were used as an input. In the present study, the previous hours $PM_{2.5}$ data were included as an input data in addition to the inputs used in the virtual monitor and forecast models.

Table 4- 7 Summary of architecture followed in developing seasonal forecast models.

Feature	Spring	Summer	Fall	Winter
No. neurons in hidden layer	10	15	10	25
Hidden layer activation function	Sine	Tanh	logistic	Gaussian complement
Output layer activation function	Sine	Gaussian complement	logistic	Logistic
No. training epochs	500	500	500	500
Training runtime (min: sec)	11:57	1:28	3:19	4.30
No. training set patterns	3356	2528	3707	3759
No. test set patterns	1119	843	1236	1253
No. production set patterns	1119	843	1235	1253
Production set R^2	0.65	0.58	0.76	0.78

Virtual monitor. The addition of persistence of data improved the predictive ability of the spring virtual monitor model. Model performance improved from an R^2 of 0.72 to 0.80. Compared to the virtual monitor without lagged data, this model was able to predict both the higher and lower values with precision. Figure 4-13 shows the prediction ability of the spring virtual monitor model with the addition of persistence of data. The summer virtual monitor model, which had poor predictive ability (0.41) before the addition of lagged data, improved its prediction with the addition of lagged data (0.65). This model was able to predict the low values (see Figure 4-14) better, compared to the virtual monitor model without lagged input data (see Figure 4-2). The predictive ability of the fall virtual monitor model improved from an R^2 value of 0.66 to 0.80 with the addition of 1-hr lagged data as an input (Figure 4-15). The winter season virtual monitor model with 1-hr lagged data was able to make better generalizations than the model without lagged data. R^2 values increased from 0.72 to 0.84 with the addition of lagged PM data (Figure 4-16).

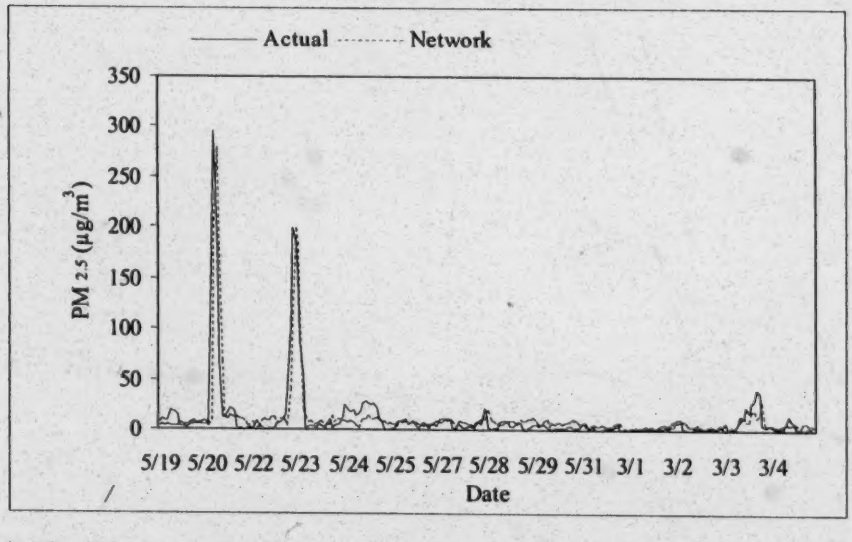


Figure 4- 13 Edmonton East Spring virtual monitor model with lagged data. $R^2 = 0.80$.

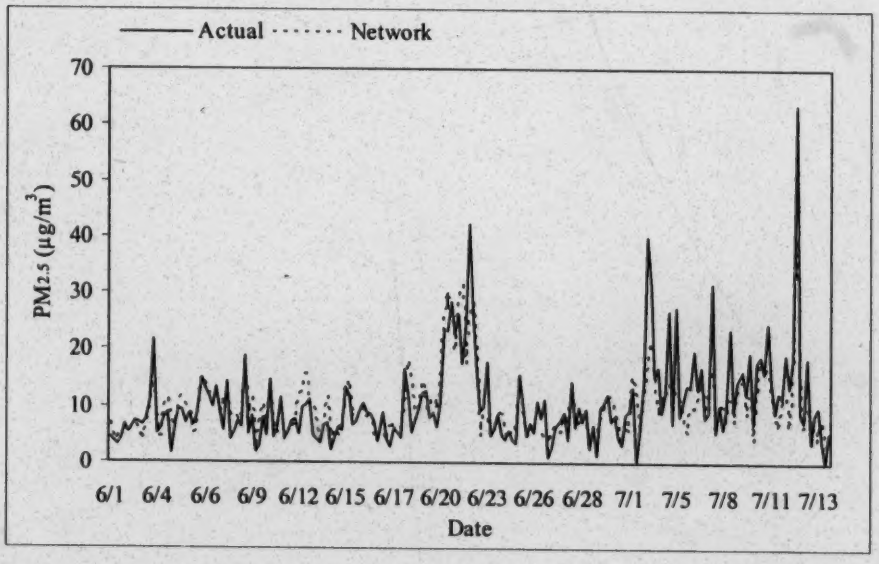


Figure 4- 14 Edmonton East summer virtual monitor model with lagged data. $R^2 = 0.65$.

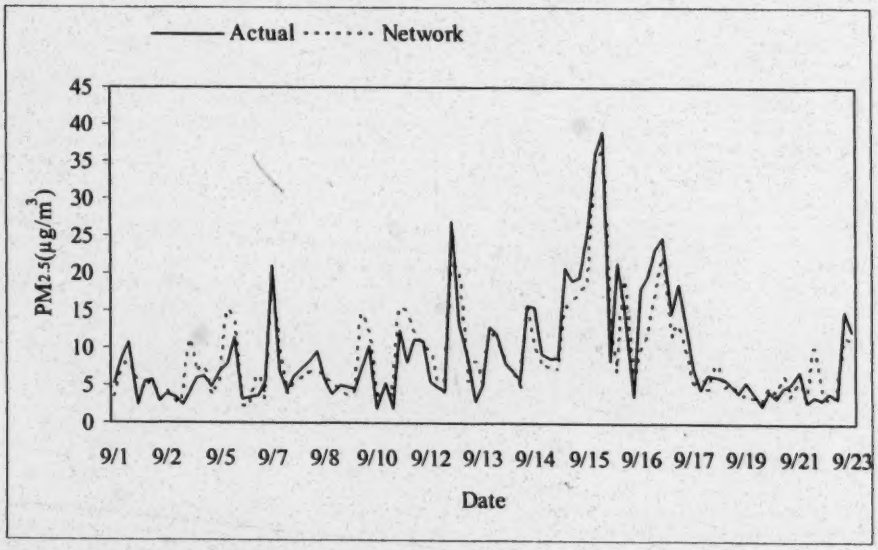


Figure 4- 15 Edmonton East fall virtual monitor model with lagged data. $R^2 = 0.80$.

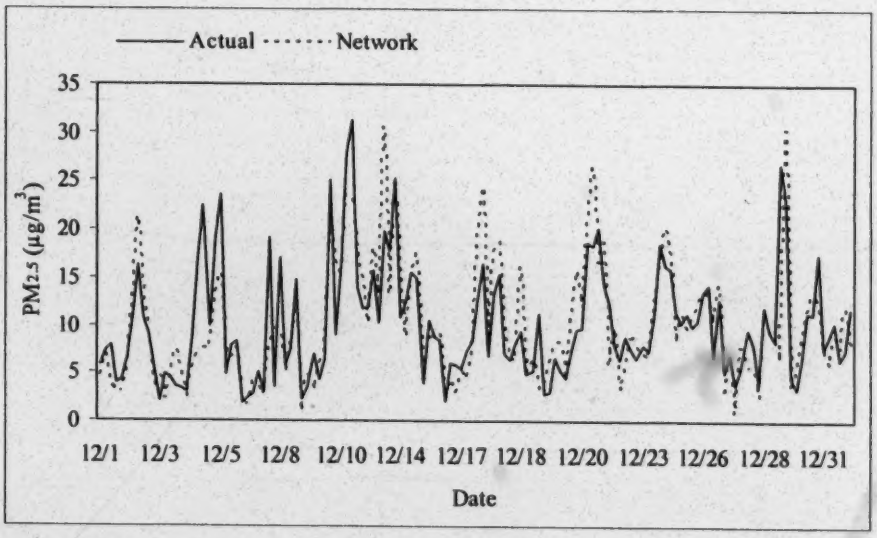


Figure 4- 16 Edmonton East winter virtual monitor model with lagged data. $R^2 = 0.84$.

Forecast model. The prediction ability of the spring forecast model improved from an R^2 value of 0.65 to 0.78 with the addition of 1-hr lagged data as an input (Figure 4-17). However, the addition of the previous hour's PM data did not improve the prediction ability of the summer season model (Figure 4-18). The R^2 value remained the same at 0.58. The fall forecast models prediction ability did not improve much as a result of the addition of 1-hr lagged data as an input (Figure 4-19). The R^2 value was 0.79, an

increase of 0.03 from the original 0.76. The addition of persistence (lagged data) of data did not improve the prediction ability of the winter forecast model (Figure 4-20). The prediction ability decreased from an R^2 of 0.78 to 0.74 when previous hour PM data was added. Further study is warranted to improve the prediction ability of forecast models. Table 4-8 summarizes the persistence model characteristics.

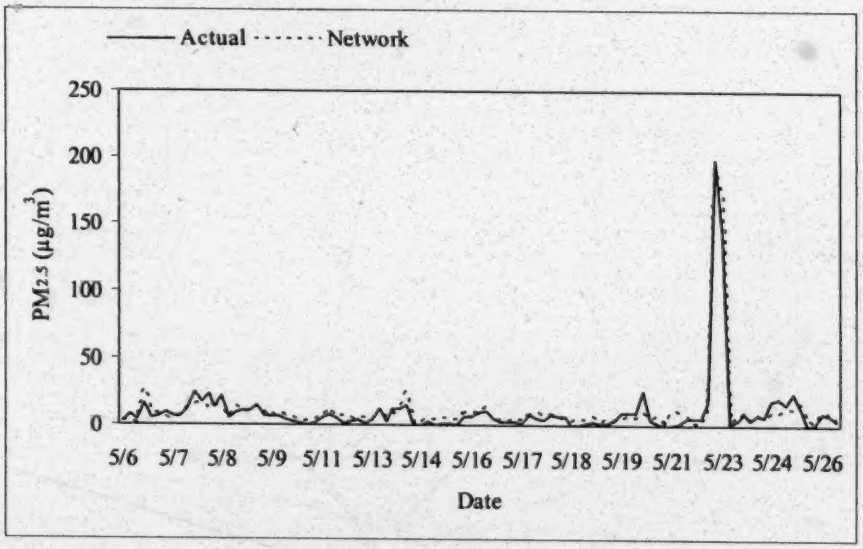


Figure 4- 17 Edmonton East Spring forecast model with lagged data. $R^2 = 0.78$.

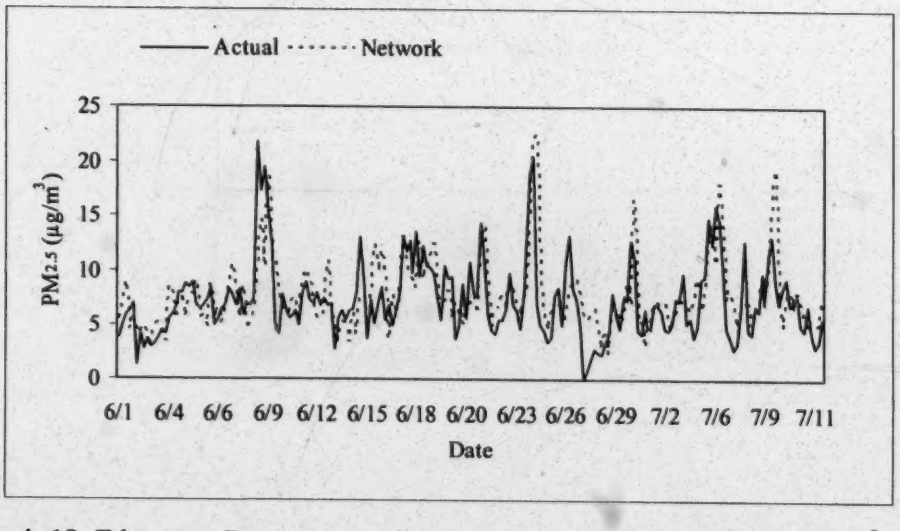


Figure 4- 18 Edmonton East summer virtual monitor model with lagged data. $R^2 = 0.58$.

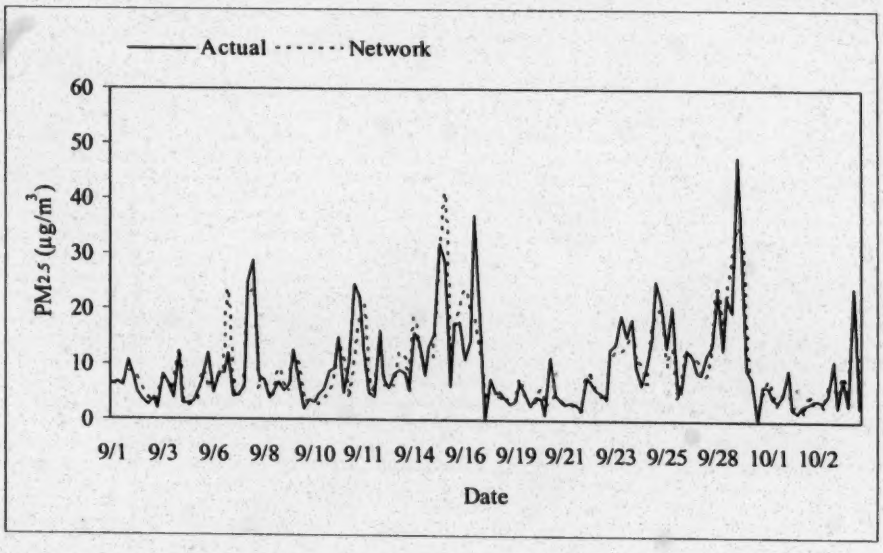


Figure 4- 19 Edmonton East Fall forecast model with lagged data. $R^2 = 0.79$.

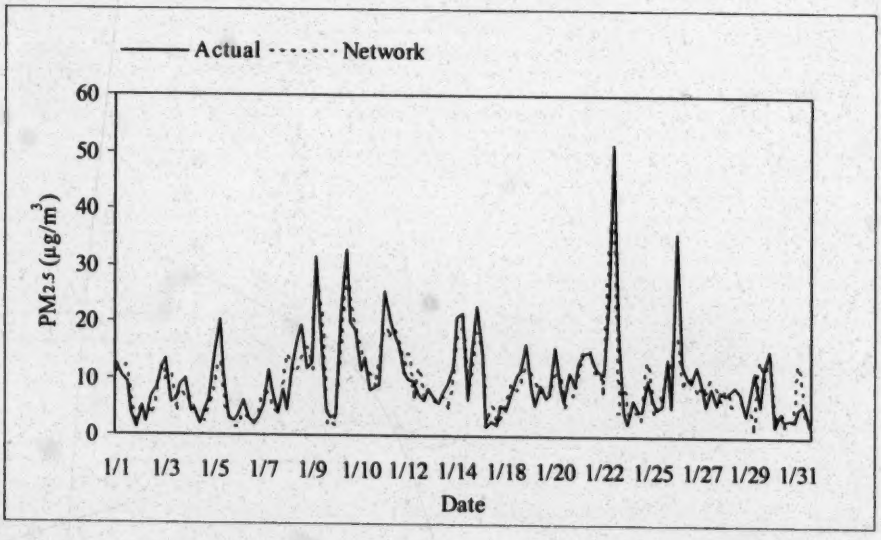


Figure 4- 20 Edmonton East Winter forecast model with lagged data. $R^2 = 0.74$.

Table 4- 8. Summary of virtual monitor and forecast models with lagged data as an input.

Model	Model characteristics	Spring	Summer	Fall	Winter
Virtual monitor	No. of hidden layer neurons	15	10	15	15
	No. of epochs	500	500	500	500
	R^2 value	0.80	0.65	0.80	0.84
Forecast	No. of hidden layer neurons	30	10	15	15
	No. of epochs	500	500	500	500
	R^2 value	0.78	0.58	0.79	0.74

4.5 Conclusions

ANNs are a promising tool for predicting and forecasting $PM_{2.5}$ in Edmonton when a separate model is developed for each season. The systematic approach followed in this research for developing the model gave better predictions compared to other models in the literature. The developed virtual monitor model will act as a standby to the real-time monitors. Also, the virtual monitor model can be used as an invaluable quality control tool. Addition of previous hour's $PM_{2.5}$ data did improve the performance of virtual monitor models. However, these models cannot be used a virtual monitor, rather they can be used as a quality control tool. The forecast model can be used to give 1-hr beforehand forecast of $PM_{2.5}$ data. Moreover, regular pollutant updates will forewarn the people of any air quality episodes. Thus the models developed in this research will assist in better monitoring and forecasting the pollutant $PM_{2.5}$ for the city of Edmonton.

In the present study, the forecast models were able to make 1-hr beforehand predictions. More studies are required to improve the forecasting window. Usage of precipitation data as an input should be attempted in the future to improve the prediction ability of the models. Attempts should be made to use dispersion modeling techniques for calculating the meteorological variables.

4.6 References

Abdul-Wahab, S.A.; Al-Alawi, S.M., 2002. Assessment and Prediction of Tropospheric Ozone Concentration Levels using Artificial Neural Network; *Environmental Modelling and Software*, **17**, 219-228.

Baawin, M.S.; Nour, M.H.; Gamal El-Din, M. Applying Artificial Neural Network Models for ENSO Prediction Using SOI and NIÑO3 as Onset Indicators; In *Proc. 8th Environmental and Sustainable Engineering Specialty Conf., 31st Annual CSCE Conf., New Brunswick, Canada, 2003.*

Baxter, C.W.; Stanley, S.J.; Zhang, Q.; Smith, D.W., 2002. Developing Artificial Neural Network models of water treatment processes: a guide for utilities; *J. Environ. Eng. Sci.*, **1**, 201-211.

Baxter, C.W.; Stanley, S.J.; Zhang, Q.; Smith, D.W. Developing Artificial Neural Network Process Models: A Guide for Drinking Water Utilities; In *6th Environmental Engineering Speciality Conference of the CSCE & 2nd Spring conference of the Geoenvironmental Division of the Canadian Geotechnical Society, London, Ontario,7-10 June 2000.*

Benkley, C.W.; Schulman, L.L., 1979. Estimating Hourly Mixing Depths from Historical Meteorological Data; *J. App. Meteorol.* **18**, 772-780.

Bogo, H.; Otero, M.; Castro, P.; Ozafran, M.J.; Kreiner, A.; Calvo, E.J.; Negri, R. M., 2003. Study of Atmospheric Particulate matter in Buenos Aires City; *Atmos. Environ.*, **37**, 1135-1147.

Burnett, R.T.; Cakmak, S.; Brook, J.R., 1998. The Effect of the Urban Ambient Air Pollution Mix on Daily Mortality Rates in 11 Canadian Cities; *Canadian Journal of Public Health*, **89**, 152-156.

Chaloulakou, A.; Kassomenos, P.; Spyrellis, N.; Demokritou, P.; Koutrakis, P., 2003. Measurement of PM_{10} and $PM_{2.5}$ Particle Concentrations in Athens, Greece; *Atmos. Environ.*, **37**, 649-660.

Chaloulakou, A.; Grivas, G.; Spyrellis, N., 2003a. Neural Network and Multiple Regression Models for PM_{10} Prediction in Athens: A Comparative Assessment; *J. Air & Waste Mange. Assoc.*, **53**, 1183-1190.

Chelani, A.B.; Rajghate, D.G.; Hasan, M.Z., 2002. Prediction of Ambient PM_{10} and Toxic Metals using Artificial Neural Networks; *J. Air & Waste Mange. Assoc.*, **52**, 805-810.

Cheng, L.; Sandhu, H.S.; Angle, R.P.; Myrick, R.H., 1998. Characteristic of Inhalable Particulate Matter in Alberta Cities; *Atmos. Environ.*, **32**, 3835-3844.

Cheng, L.; McDonald, K.M.; Angle, R.P.; Sandhu, H.S., 1998a. Forest Fire Enhanced Photochemical Air Pollution- A Case Study; *Atmos. Environ.*, **32**, 673-681.

El-Din, A.G.; Smith, D.W., 2002. A Neural Network Model to Predict the Wastewater Inflow Incorporating Rainfall Events; *Water Res.*, **36**, 1115-1126.

Environment Canada, 1999. Federal-Provincial Working Group on Air Quality Objectives and Guidelines. *National Ambient Air Quality Objectives for Particulate Matter: Science Assessment;* 1999: Environment Canada and Health Canada.

Environmental Protection Agency (EPA), *Guideline on Air Quality Models (Revised).* Office of airquality planning and standards, Research triangle Park, NC 27711, July1986.

Fernandez de Castro, B.M.; Sanchez, J.M.P.; Manteiga, W.G.; Bande, M.F.; Bermudez Cela, J.R.; Fernandez, J.J.H., 2003. Prediction of SO_2 Levels using Neural Networks; *J. Air & Waste Mange. Assoc* , **53**, 532-539.

Gardner, M.W.; Dorling, S.R., 1999. Modeling and Prediction of Hourly NO_x and NO_2 Concentrations in Urban Air in London; *Atmos. Environ*, **33**, 709-719.

Gardner, M.W.; Dorling, S.R., 1998. Artificial Neural Networks (The Multilayer Perceptron) - A Review of Applications in the Atmospheric Sciences; *Atmos. Environ.* , **32**, 2627-2636.

Harrison, R. M.; Deacon, A.R.; Jones, M.R., 1997. Sources and Processes Affecting Concentrations of PM_{10} and $PM_{2.5}$ Particulate Matter in Birmingham (U.K.); *Atmos. Environ.*, **31**, 4103-4117.

Hasham, F.A.; Kindzierski, W. B.; Stanley, S. J., 2004. Modeling of Hourly NO_x Concentrations using Artificial Neural Networks; *J. Environ. Eng. Sci*, **3**, 111-119.

Jain, A.K.; Mao, J., 1996. Artificial Neural Network: A Tutorial; *Computer*, **29**, 31-44.

McCullum, K.; Kindzierski, W.B.; Gamal El-Din, M.; Myrick, B. Elemental Trends in Urban Ambient Particulate Matter in Downtown Edmonton and Calgary, Alberta; In *Cold Regions Engineering and Construction Conference Proceedings, Edmonton, Alberta, Canada, 2004*. Paper # 57.

McCullum, K.; Hasham, F.; Kindzierski, W.B. Ambient Air Quality Trends using Data collected throughout Alberta; In *proceedings of A&WMA's 97th Annual Conference and Exhibition, Indianapolis, Indiana, 2004a*; A&WMA: Pittsburg, PA, 2004; Paper AA-454.

McDonald, K.; Shepherd, M., 2004. Characterization of Visibility Impacts Related to Fine Particulate Matter in Canada; *J. Air & Waste Mange. Assoc.*, **54**, 1061–1068.

McKendry, I.G., 2002. Evaluation of Artificial Neural Networks for Fine Particulate Pollution (PM_{10} and $PM_{2.5}$) Forecasting; *J. Air & Waste Mange. Assoc.*, **52**, 1096-1101.

McKendry, I.G.; Hacker, J.P.; Stull, R., Sakiyama, S.; Mignacca, D.; Reid, K., 2001. Long-range Transport of Asian Dust to the Lower Fraser Valley, British Columbia, Canada; *J. of Geophys. Res.-atms.*,**106**, 18361-18370.

Melas, D.; Kioutsioukis, I.; Ziomas, I.C., 2002. Neural Network Model for Predicting Peak Photochemical Pollutant Level; *J. Air & Waste Manage. Assoc.*, **50**, 495-501.

Mitchelle, A.E.J.; Timbre, K.O. Atmospheric Stability Class from Horizontal Wind Fluctuation. In *72nd Annual meeting of air pollution control Association. Cincinnati, OH, June 24-29, 1979.*

Myrick, R.H.; Sakiyama, S.K.; Angle, R.P.; Sandhu, H.S., 1994. Seasonal Mixing Heights and Inversions at Edmonton, Alberta; *Atmos. Environ.*, **28**, 723-729.

Perez, P.; Palacios, R.; Castillo, A., 2004. Carbon Monoxide Concentration Forecasting in Santiago, Chile; *J. Air & Waste Mange. Assoc.*, **54**, 908-913.

Perez, P.; Reyes, J., 2001. Prediction of Particulate Air Pollution using Neural Techniques; *Neural Computing and Applications,* **10**, 165-171.

Sandhu, H.S. *Ambient Particulate Matter in Alberta*; Report prepared for Science and Technology Branch, Alberta Environmental Protection, No. 1494-A-9805, Edmonton, Alberta, 1998.

Su, D.; Gamal El-Din, A.; Gamal El-Din, M.; Idriss, A.; Wiens, B. Historical $PM_{2.5}$ Monitoring Data Trends in Edmonton and Calgary: Investigation of their Potential Use for $PM_{2.5}$ Modeling; In *Cold Regions Engineering and Construction Conference Proceedings, Edmonton, Alberta, Canada, 2004*; Paper # 55.

Su, D.; Gamal El-Din, A.; Gamal El-Din, M.; Idriss, A.; Wiens, B. Historical Ozone Monitoring Data Trends in Edmonton and Calgary: Investigation of their potential use for

Ozone Modelling; In *Cold Regions Engineering and Construction Conference Proceedings, Edmonton, Alberta, Canada, 2004a*; Paper # 56.

Su, D.; Gamal El-Din, M.; Selvaraj, M. *Modelling $PM_{2.5}$ in Calgary and Edmonton: Application of Artificial Neural Networks*; Report prepared for Air Section and Science Standards Branch, Alberta Environment, Edmonton, Alberta, 2004b.

Weber, R. O., 1997. Estimators for the Standard Deviation of Horizontal Wind Direction; *J. Appl. Meteorol.*, **36**, 1403-1415.

5.0 CALGARY EAST $PM_{2.5}$ MODELS

5.1 Introduction

Air borne particulate matter smaller than 2.5μm is called $PM_{2.5}$ (NRC, 1998). $PM_{2.5}$ is considered toxic as defined in the section 64 of Canadian Environmental Protection Act, 1999 (Environment Canada, 2000). Increase in $PM_{2.5}$ concentration is linked to morbidity and mortality in the general population (Schwartz et al., 1996; Sunyer et al., 1996). Combustion related emissions from vehicles, industries and domestic heating are the some of the anthropogenic sources of $PM_{2.5}$ in addition to the naturally emitted sources.

Receptor modeling is one way of monitoring and understanding $PM_{2.5}$, their emission and sources. Some of the common receptor modeling techniques that is used in the field of air monitoring are multi-variate analysis (Hopke et al., 1976), back trajectory (Ashbaugh, 1983), neural networks (Gardner and Dorling, 1998) and time series (Hies et al., 2000). Due to the complex and non-linear relationship between $PM_{2.5}$ and meteorology, neural networks turned out to be a best option for their modeling.

Artificial Neural Networks (ANNs), an Artificial Intelligence technique predicts the future trends based on past historical data patterns. ANNs was used in the past to predict pollutants such as SO_2 (Fernandez et al., 2003), PM_{10} (Chaloulakou et al., 2003a; Chelani et al., 2002), O_3 (Abdul-Wahab and Alawi, 2002; Gardner and Dorling, 2001), NO_x (Gardner and Dorling, 1999; Hasham et al., 2004) and $PM_{2.5}$ (McKendry, 2002; Perez et al., 2000). Application of ANNs for prediction of $PM_{2.5}$ was not successful due to the complexity in the formation of $PM_{2.5}$.

The main objective of this research is to develop models for predicting and forecasting the concentration of $PM_{2.5}$ for the city of Calgary, Canada using a systematic approach. The predictive model can be used as a virtual monitor in case of failure to the

---

A version of this chapter will be submitted to the Journal of Applied Meteorology.

instrument. On the other hand, the forecast model can make 1-hr beforehand $PM_{2.5}$ values and forewarn people of any pollution episodes. The other objective of this research is to investigate the input variables that contribute towards the concentration of $PM_{2.5}$.

5.2 Data and Methods

5.2.1 Study area

Calgary, Southern city in the province of Alberta, Canada is located on the foothills of the Rocky Mountains. Calgary weather is characterized by long winters and short summers. Chinook, a warm, dry, gusty wind from mountains accompanied by rapid temperature changes is one of the characteristics of Calgary winters (Nkmidirim, 1988; Angle and Sakiyama, 1991). Sources of particulate matter in Calgary include transportation, soil, road dust, wood burning, steel foundries, and cement distribution centres (Cheng et al., 1998). Air quality is monitored on a continuous basis by three monitoring stations: East, Central and North-west located in and around the City. Some of the air quality parameters that are monitored in these stations are Carbon monoxide (CO), Hydrogen Sulphide (H_2S), Nitric oxide (NO), Nitrogen dioxide (NO_2), Sulphur dioxide (SO_2) and Total Hydrocarbons (THC). In addition, wind direction, wind direction deviation and wind speed are the meteorological parameters that are monitored in these stations. However, relative humidity and temperature are monitored only in the North-west station.

5.2.2 Artificial Neural Network

Artificial Neural Networks (ANNs), an Artificial Intelligence (AI) technique is similar to the working of human brain. ANNs have been widely used in various fields for pattern classification and prediction (El-Din and Smith, 2002). One of the advantages of using ANN is that no priori knowledge of the study area is required (Gardner and Dorling, 1998). The most common ANNs technique is the Multi Layer Perceptron (MLP). A typical MLP (Figure 1) consists of three layers: input, hidden and output. Each layer consists of processing elements called neurons, which passes information between different layers. A neuron consists of an activation function which stimulates the output

to the next layer. The MLP learns through a supervised learning methodology called back propagation technique, which reduces the error between the predicted and original values. Data patterns which consist of a set of inputs and its related output is used in training the models to predict the future trends. The data patterns are normally divided in to three categories: training, testing and production. The training data set is

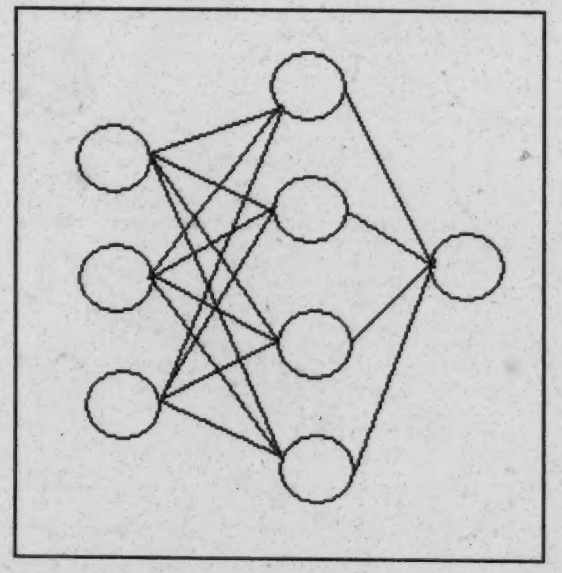


Figure 5- 1 A three-layered neural network

used to train the model in order to generalize the future trends. The test set is usually used to avoid memorization during training of the model. The production data set which is independent of both training and test data set is finally used to ascertain the prediction capability of the trained network. Training of the model is done by passing the entire set of data pattern through the network (epochs). Evaluating the number of epochs and number of hidden layer neurons are some of the problems encountered by researchers when using ANNs. Several performance criteria such as Coefficient of multiple determinations (R^2), correlation coefficient (r) were used to check the generalization ability of the developed network (Su, 2004).

5.2.3 Methodology

In the present study, historical data was obtained from the Calgary East station. Hourly averaged pollutant parameters such as CO, NO, NO_2, SO_2, THC and O_3 were used as input. Wind speed, wind direction, mixing height, temperature, opacity, and relative humidity are some of the meteorological parameters that were used as input for the model development. Mixing height was calculated using a theoretical study done by Benkley and Schulman (1979). Indexed input variables for each month of the year (12), day of the week (7) and hour of the day (24) was also included, to simulate the temporal variations observed in the values of meteorological and pollutant parameters. For example, if the pollutant concentration was observed on 12^{th} (Thursday) November at 14hrs, then an Indexing of 1 for November, Thursday and 14hrs and 0 for other months, days and hours was used. Table 5-1 shows the parameters used in model development. The

Neuroshell-2 (version 4.0) from Ward systems Inc., neural network software was used in developing the models. The models were developed using a systematic approach outlined by El-Din and Smith (2002). The data patterns were divided in to four seasons: Fall (September, October, November), Winter (December, January, February), Spring (March, April, May) and Summer (June, July, August) due to the success of a similar study done in Edmonton, Canada (Su et al., 2004). Two kinds of model were developed for all the four seasons, prediction and forecast. The performances of the models were ascertained by looking at the Coefficient of multiple determination value (R^2).

The following is the systematic approach that was used in developing the models

1. Screen the historical data pattern for erroneous or missing data using Microsoft Excel.
2. Divide the screened historical data into training, testing and production data set in a ratio of 3:1:1.

Table 5- 1 Parameters used in developing models for different seasons.

Parameter	Month	Day	Hour	Pollutant	Meteorology
Input	January	Sunday	0	CO	Mixing height
	February	Monday	1	NO	Wind Direction
	March	Tuesday	2	NO_2	Wind Speed
	April	Wednesday	3	SO_2	Temperature
	May	Thursday	4	O_3	Relative Humidity
	June	Friday	5	THC	Opacity
	July	Saturday	6		
	August		7		
	September		8		
	October		9		
	November		10		
	December		11		
			12		
			13		
			14		
			15		
			16		
			17		
			18		
			19		
			20		
			21		
			22		
			23		
Output				$PM_{2.5}$	

3. Determine the best activation function combination for the hidden and output layer by varying them at three different settings of number of hidden layers and number of epochs (Table 5-2). The activation function combination that was used in previous research (Su et al., 2004) in Edmonton was used in this research for developing the models. Table5-3 shows the activation function used in the model development.

Table 5- 2 Number of hidden layer neurons and training epochs settings for evaluating hidden and output layer activation functions.

Setting	No. of neurons in hidden layer	No. of training epochs
Low	5	1000
Middle	20	4000
High	40	7500

4. Optimize the best activation function combination by varying the number of epochs and number of hidden layer neurons.
5. Check for persistence by the addition of previous hour pollutant data ($PM_{2.5}$) as an input.
6. Determine the input variables that are affecting the output variable. This is done by removing the input variables one at a time and retraining the model with same network architecture that was used before for developing the model. Any change in R^2 value by 0.03 is due to the importance of that variable in final prediction.

5.3 Results and Discussion

5.3.1 Prediction model

A thorough data analysis was carried out and all the erroneous data was removed. The number of data patterns available for model development after data analysis was 2225, 2700, 2150 and 2080 for fall, winter, spring and summer season respectively. Based on the systematic approach, season specific models were developed using activation functions shown in Table 5-3.

Table 5- 3 Type of activation function that is used in model development for different seasons based on previous study (Su et al., 2004).

Season	Hidden layer activation function	Output layer activation function
Fall	Logistic	Logistic
Winter	Gaussian complement	Logistic
Spring	Sine	Sine
Summer	Tanh	Gaussian complement

5.3.1.1 Spring model

Figure 5-2 shows the plot for spring season through which the best model architecture was found out. From the Figure-2 the best model architecture was found to be 20 hidden layer neurons and 500 epochs with the R^2 value of 0.61 based on the systematic approach. The spring model was able to follow the lower values of historical data trends well, but slightly over predicted the higher values. Figure 5-3 shows the prediction ability Calgary spring prediction model. The Calgary spring model had lesser performance compared to the Edmonton spring model.

Figure B-1 (Appendix-C) shows the plot of network concentration versus actual concentration. In this figure, the points are distributed evenly on both sides of the ideal case line (black line), which clearly depicts good prediction ability. Also, residual analyzes shown in Figure B-2 did not show any discernible trends.

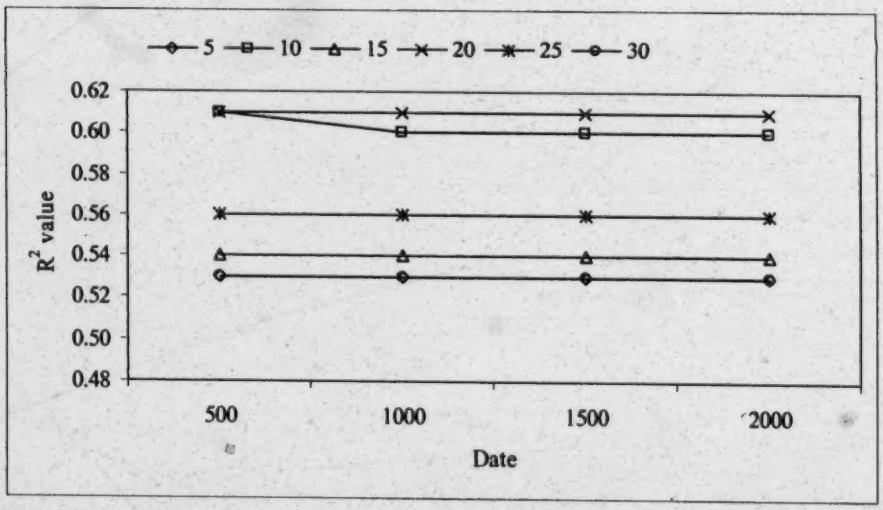


Figure 5- 2 Optimization of model architecture for spring season virtual monitor model.

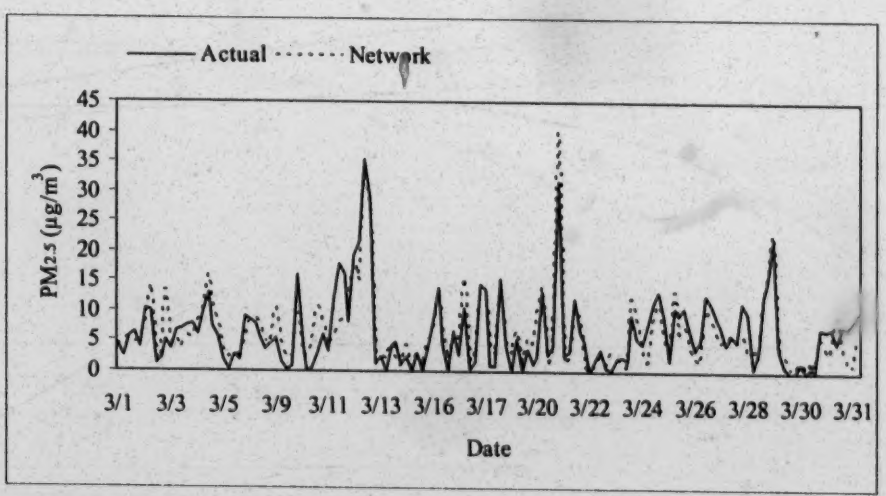


Figure 5- 3 Calgary East spring prediction model. $R^2 = 0.61$.

5.3.1.2 Summer model

Figure-5-4 shows the optimization of summer prediction model. Accordingly, it is evident from the Figure that 15 hidden layer neurons and 500 training epochs give optimum network performance of 0.71. Despite the limited data history available, the network was able to predict the highest peak in the production data set. However, the ANN tended to over-predict small peaks and under-predict mid-level peaks. Figure 5-5

shows the prediction ability of the summer prediction model. This model's performance was significantly better than the summer models for Edmonton.

Network prediction for the summer season prediction model is in good agreement with the actual concentration which is evident from the Figure B-3. Residual analysis shows no obvious trends (Figure B-4).

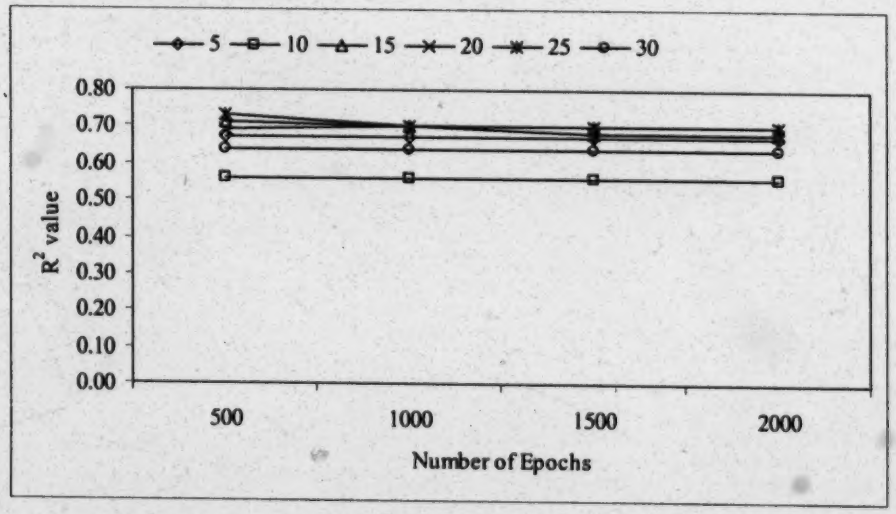


Figure 5- 4 Optimization of model architecture for summer season prediction model.

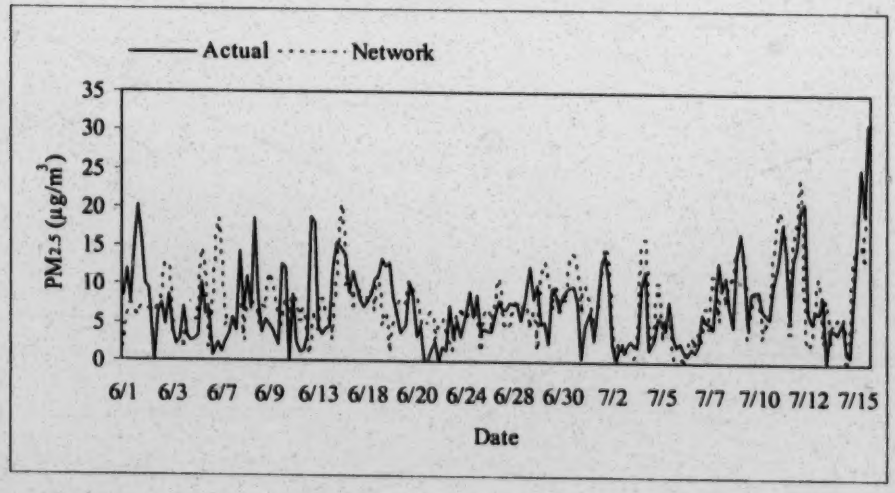


Figure 5- 5 Calgary East summer prediction model. $R^2 = 0.71$

5.3.1.3 *Fall model*

According to the systematic approach, 15 hidden neurons, 500 epochs resulted in a R^2 value of 0.72. Figure 5-6 shows the optimization of model characteristics for Calgary prediction model for fall season. The Calgary East fall prediction model's performance is shown in Figure 5-7 for November 2002 data. The performance of the network when applied to the production set of data was at a R^2 level of 0.72. The model generally followed the $PM_{2.5}$ concentration trends well, with no consistent under- or over-prediction of peak levels. The training time of the model was 1:11 minutes. The model used a logistic activation function in both the hidden and output layers.

Figure B-5 shows the plot of Network concentration versus the Actual concentration. Except few points, most of the points were evenly distributed on either side of the solid line. Thus, the network shows good prediction agreement with the actual concentration. The residual analysis did not show any discernible trends (Figure B-6).

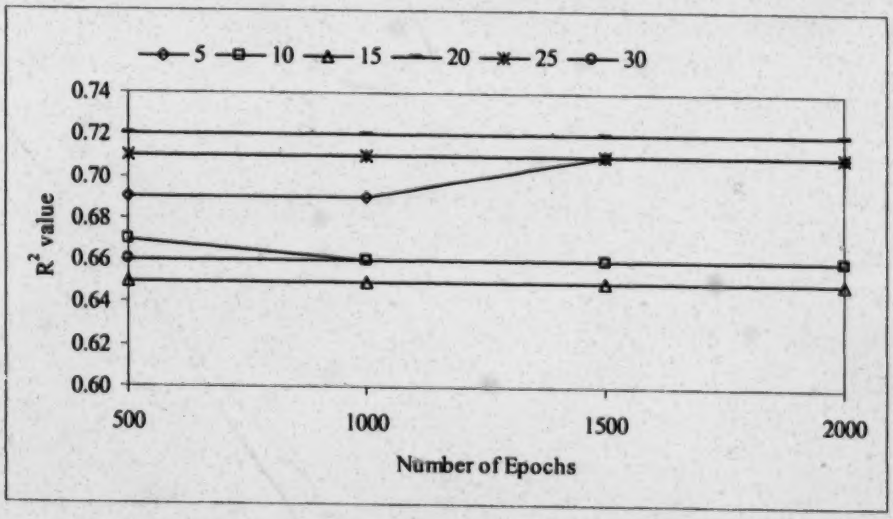


Figure 5- 6 Optimization of model architecture for fall season prediction model.

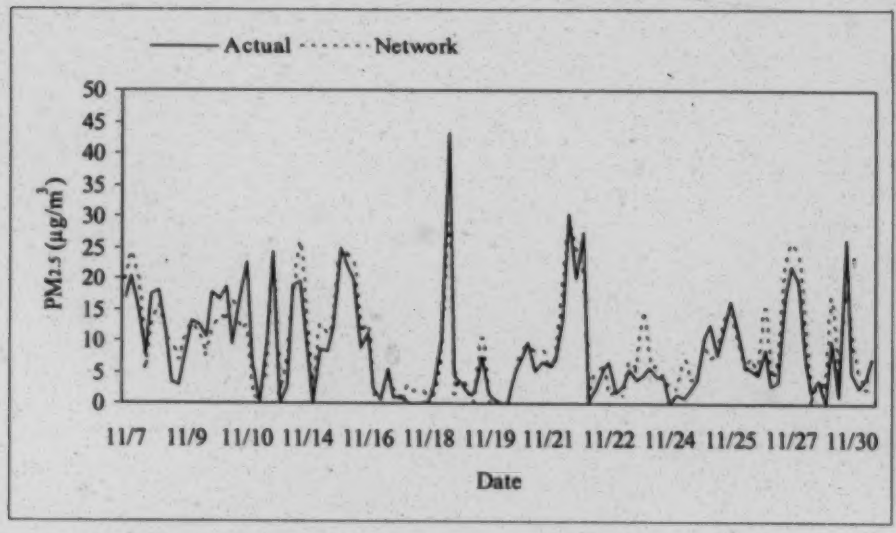


Figure 5- 7 Calgary East Fall prediction model. $R^2 = 0.72$.

5.3.1.4 Winter model

The Calgary winter prediction model was least successful at modelling $PM_{2.5}$ concentrations, generating a production set R^2 of 0.46, and may be a reflection of Calgary's erratic winter climate. The model was built with 10 hidden layer neurons, requiring 500 epochs and 0:37 minutes to train. The performance for December 2003 data is shown in Figure 5-9. The ANN struggled with modelling peak levels and low points, under-predicting most high $PM_{2.5}$ concentrations. Unlike the other ANN models, the Calgary winter model also had difficulty modelling the dips (valleys) in the $PM_{2.5}$ concentration. Figures B-7 and B-8 supports the poor prediction ability of the winter model which is evident from the scattered distribution of points.

The reason for winter seasons poor prediction ability may be due to the prevalence of Chinook wind, a condition in which a sudden increase in temperature causes melting of the ice cover. This unusual weather pattern is due to the western disturbance from the Rocky Mountains, which creates a pressure gradient from the top of the mountains (Nkmidirim, 1988; Angle and Sakiyama, 1991). Lack of wind direction deviation data, which was used as a surrogate of stability conditions (Hasham et al., 2004) may be the other reason for the winter seasons poor prediction ability. According to McCullum et al.

(2004) elemental and ionic compounds such as chloride ion, sodium, sodium ion and nitrate ion are found in Calgary during the winter and spring months due to the re-suspension of road salts. Non-inclusion of these parameters might be the reason for networks mediocre ability during the seasons.

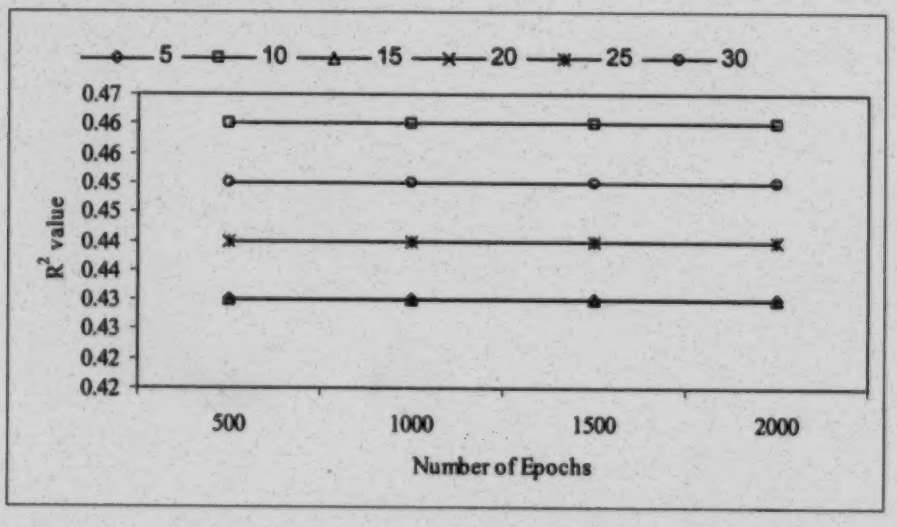


Figure 5- 8 Optimization of model architecture for winter season prediction model.

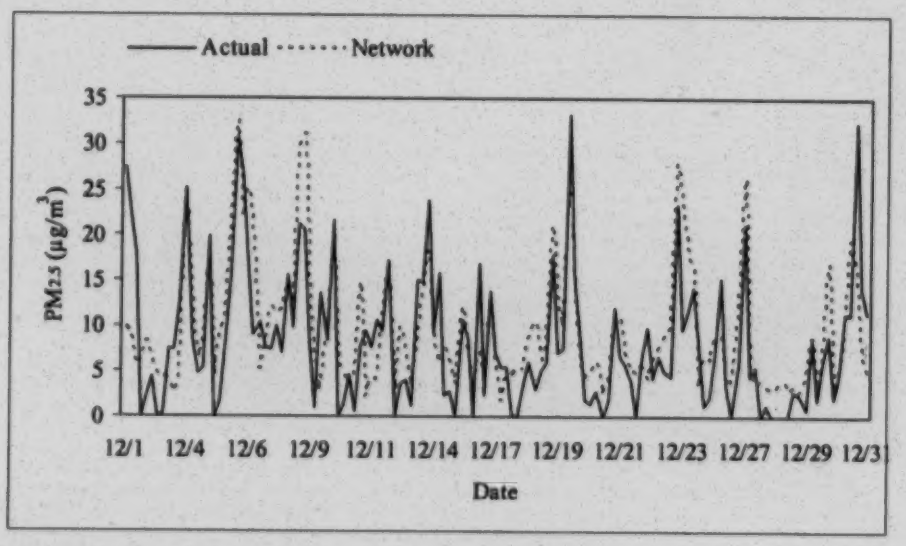


Figure 5- 9 Calgary East winter prediction model. $R^2 = 0.46$.

5.3.2 Prediction model with lagged data as an input

Addition of previous hours $PM_{2.5}$ concentration as an input improved the prediction ability of seasonal models.

5.3.2.1 Spring model

The predictive ability of the spring season model improved from 0.61 to 0.76 with the addition of lagged input data as an input. This model was trained with 500 training epochs and 10 hidden layer neurons. Figure 5-10 shows the network prediction ability for the spring prediction model when previous hour's data was included. The network was able to follow the dips and valleys of the historical data with precision.

Figure B-9 shows the plot of Network versus Actual $PM_{2.5}$ concentration. From the Figure, it is clearly visible that most of the points were evenly distributed on either side of the solid line, thus confirming good prediction ability of the model. Also, residual analyses performed did not show any unusual patterns or trends (Figure B-10).

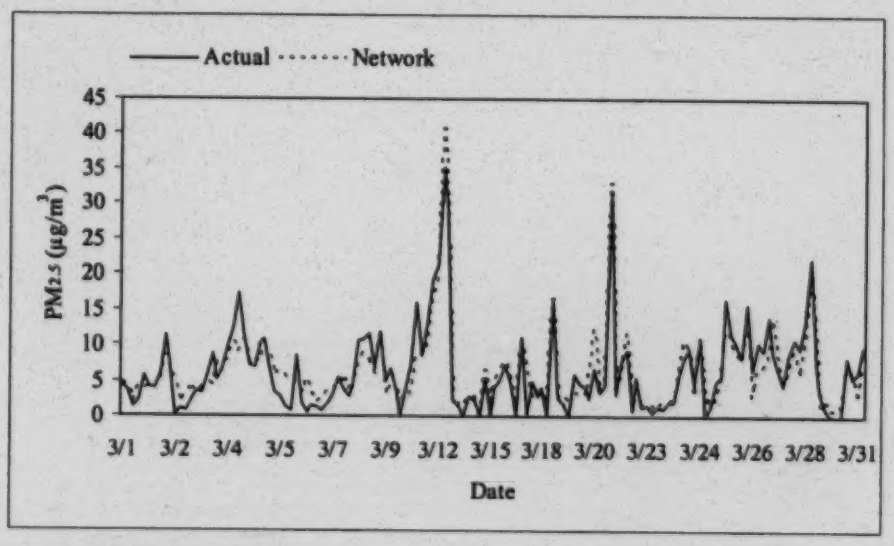


Figure 5- 10 Calgary East Spring prediction model with lagged data. $R^2 = 0.76$.

Summer model

Addition of lagged data as an input, improved the prediction ability of summer season model marginally (0.71 to 0.74). The model was not able to follow the higher and lower historical trends precisely (Figure 5-11). But the model was able to follow the trends to some extent.

The plot of the Network versus Actual concentration is shown in Figure B-11. In that Figure, the points were distributed equally around the ideal case line (solid line). Residual analyzes (Figure B-12) did not show any unusual pattern or trends.

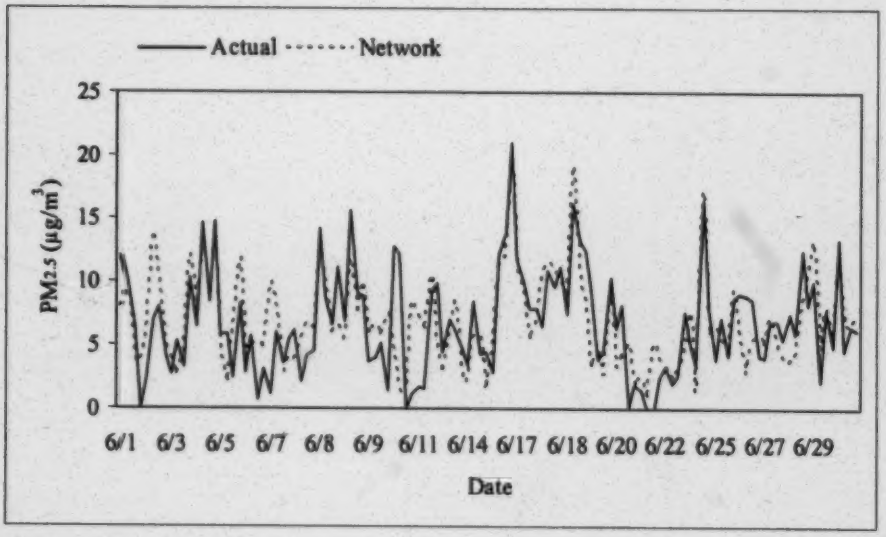


Figure 5- 11 Calgary East Summer prediction model with lagged data. $R^2 = 0.74$.

5.3.2.2 Fall season

The fall season which was developed with 15 hidden layer neurons and 500 training epochs, had the best prediction of all the seasonal models. The R^2 value improved from 0.72 to 0.85 with the addition of previous hour's data. Figure 5-12 shows the networks prediction ability for fall season model. This model was able to predict the peaks and lows of $PM_{2.5}$ concentration better the model without lagged data as an input.

The plot of Network versus Actual concentration shows even distribution of points on either side of the black line (Figure B-15). Moreover, there are no discernible trends when residual analyses performed (Figure B-16).

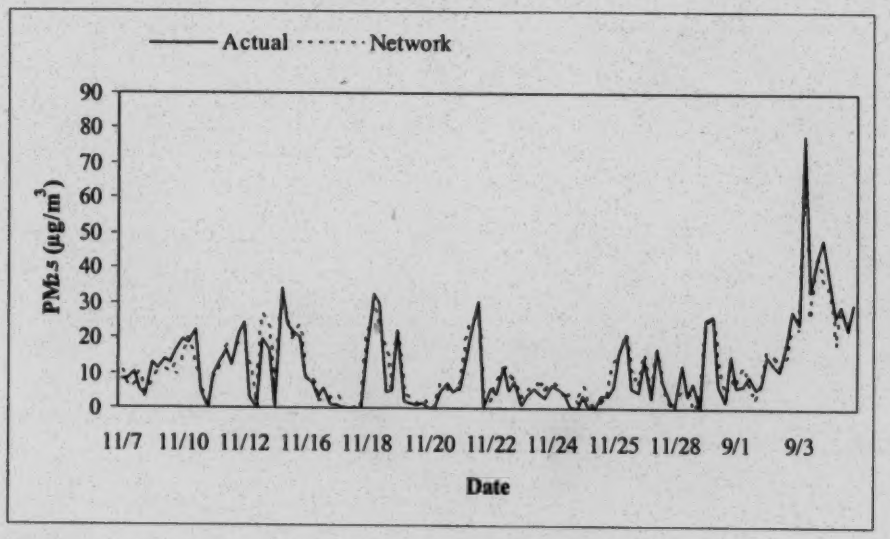


Figure 5- 12 Calgary East fall prediction model with lagged data. $R^2 = 0.85$.

5.3.2.3 Winter season

The winter season prediction model did not show any improvement in the prediction ability. The R^2 value improved from 0.46 to 0.50 with the addition of lagged data. Also, the network was not able to follow the actual trends of $PM_{2.5}$ concentration (Figure 5-13). Figure B-15 shows the Network versus Actual $PM_{2.5}$ concentration. From the Figure, it is clearly evident that the point unevenly scattered on both sides of the black line. Thus the developed winter model does not like robust one. Also, the residual analyzes (B-16) showed an unusual pattern. Table 5-4 summarizes the seasonal prediction model architecture characteristics with and without lagged data as an input.

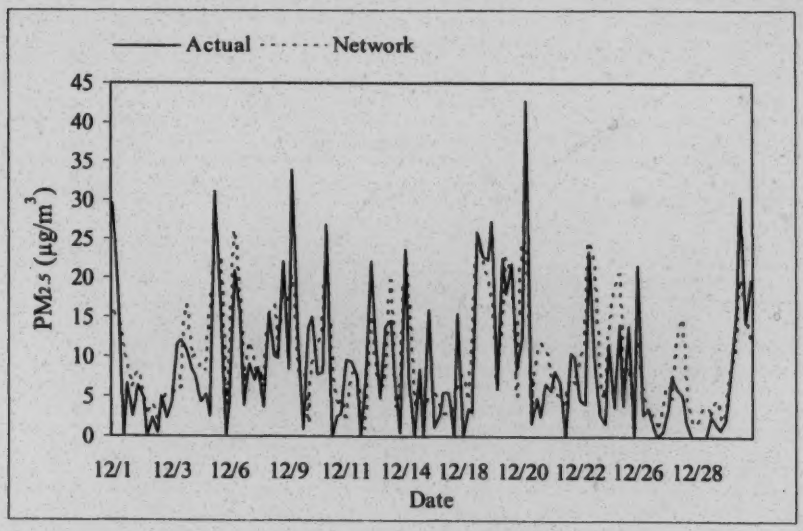


Figure 5- 13 Calgary East winter prediction model with lagged data. $R^2 = 0.50$.

Table 5- 4 Summary of virtual monitor model characteristics with and without lagged data as an input.

Architecture	Fall	Winter	Spring	Summer
Prediction model				
No. of hidden layer neurons	20	10	20	15
No. of training epochs	500	500	500	500
Coefficient of multiple determination (R^2)	0.72	0.46	0.61	0.71
Prediction model with lagged data				
No. of hidden layer neurons	15	20	10	15
No. of training epochs	200	200	200	300
Coefficient of multiple determination (R^2)	0.85	0.50	0.76	0.74

5.3.3 Forecast model

The forecast models for all the seasons were developed by including present hours $PM_{2.5}$ concentration as an input to forecast $PM_{2.5}$ concentration 1-hr in advance. All the forecast models were developed in the same way as the prediction models. This approach

recognizes the time series effects in ambient air pollution, and makes use of all available inputs for the forecast.

5.3.3.1 Spring model

The Edmonton spring forecast model architecture consists of 5 hidden layer neurons. The network required 200 epochs and 11:57 minutes to train. The resulting network performed at a R^2 of 0.67 for the production set of data. The network performance is shown in Figure 5-14. For large peaks, the network tended to under-predict $PM_{2.5}$ concentrations. Small and mid-level peaks were generally overshot.
Figure B-17 shows the plot of Network concentration versus Actual concentration. From the Figure it can be seen that the distribution of the markers is reasonably ok. Also, the residual analyses show the similar performance trend (Figure B-18).

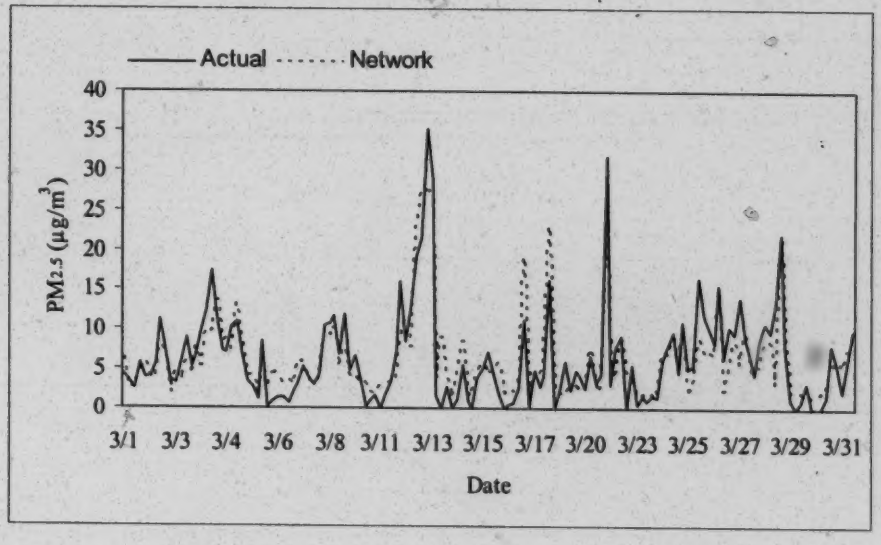


Figure 5- 14 Calgary East spring forecast model with lagged data. $R^2 = 0.67$.

5.3.3.2 Summer model

The summer forecast model performed better than the virtual monitor model, producing a R^2 of 0.68 for the production set. The ANN was able to predict the general $PM_{2.5}$ trend, but over-predicted some mid-level peaks and under-predicted high peaks The model's

architecture was found to be 5 hidden layer neurons and 300 epochs. Its performance for June 2002 data is presented in Figure 5-15.

Figure B-19 and B-20 shows the Network versus Actual concentration and residual analyzes respectively. From the Figures, the model appears to be stable.

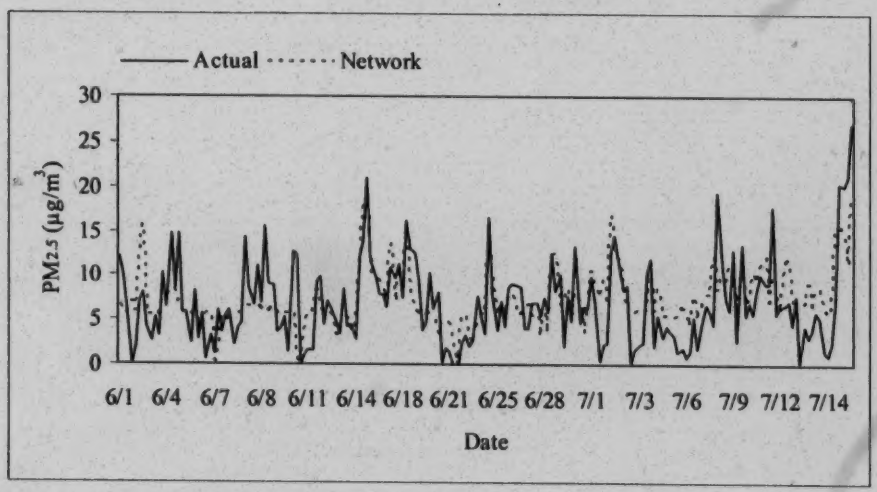


Figure 5- 15 Calgary East summer forecast model. $R^2 = 0.68$.

5.3.3.3 Fall model

The fall model performed at an equivalent level to the spring forecast, with a R^2 of 0.71 for the production data set. The fall forecast model contained 10 neurons in the hidden layer, and required 500 epochs to train. The total runtime needed for training the model was 3:19 minutes. The fall forecast model performance is shown in Figure 5-16. The figure shows that the fall model predicted the $PM_{2.5}$ concentration trends well, with no large over- or under-predictions.

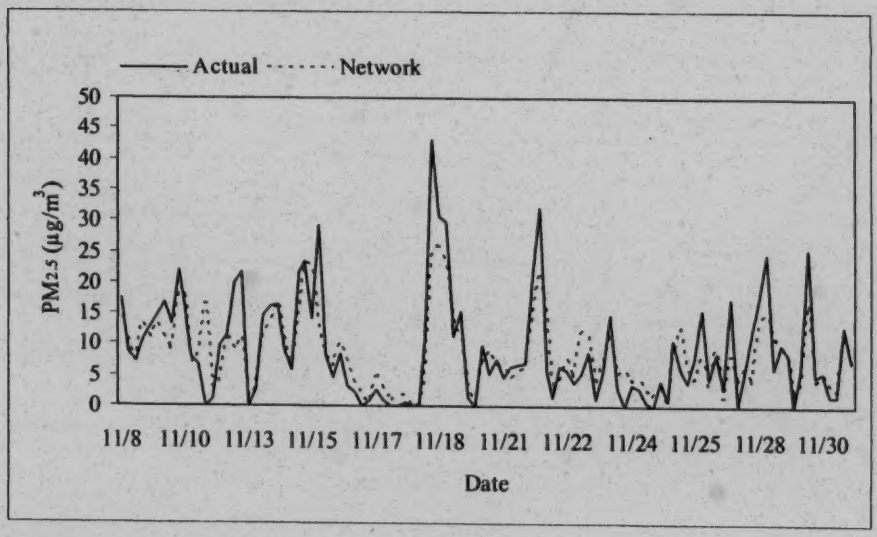


Figure 5- 16 Calgary East fall forecast model $R^2 = 0.71$.

5.3.3.4 Winter model

The Calgary winter forecast model required 15 neurons in the hidden layer. When trained in 500 epochs, the model produced a poor R^2 of 0.38, despite having the largest number of data patterns available for training of all the seasons. As seen in Figure 5-17, the model produced false positives (predicting peaks where none occurred) and missed some peak levels.

Figure B-23, shows that the markers were distributed unevenly on both sides of the solid line. Moreover, residual analyzes did show some unusual patterns (Figure B-24). Thus the model appears to be unstable.

The forecast model performance was poorer compared to the prediction model for all seasons except the spring season. The reason for which is unknown.

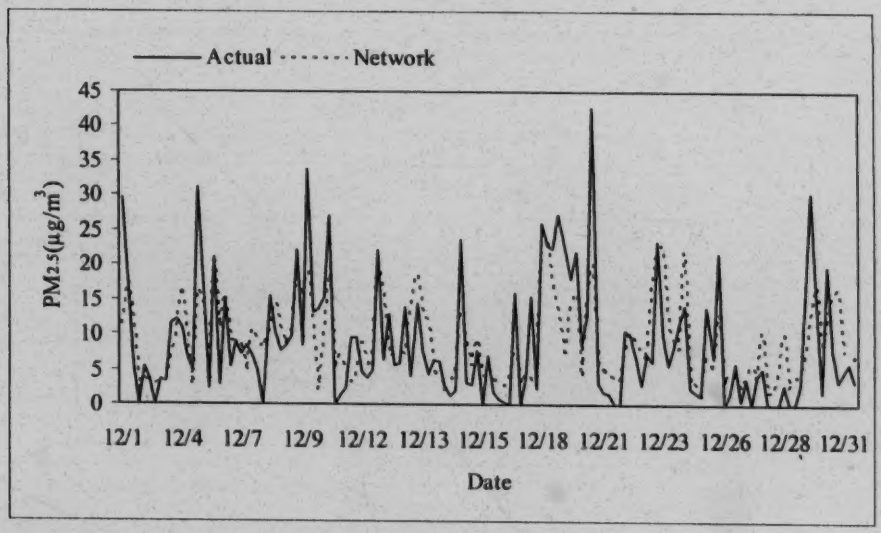


Figure 5- 17 Calgary East winter forecast model $R^2 = 0.46$.

5.3.4 Forecast model with lagged data as an input

The forecast model with lagged data as an input includes present hours and previous hours $PM_{2.5}$ as an input. All the models were developed in the same fashion as the prediction models. Residual analyzes of all the seasonal models are shown in the Appendix-C.

5.3.4.1 Spring model

The spring forecast model with lagged data as an input had an R^2- value of 0.63, when developed with 500 epochs and 5 hidden layer neurons. The model was able to predict the lower historical trends better than the higher ones. The spring forecast model without lagged data as an input performed better than the model developed with lagged data. R^2- value decreased from 0.67 to 0.63. This is contrary to the prediction model which showed increased performance with the inclusion of lagged data as an input (0.61 to 0.76).

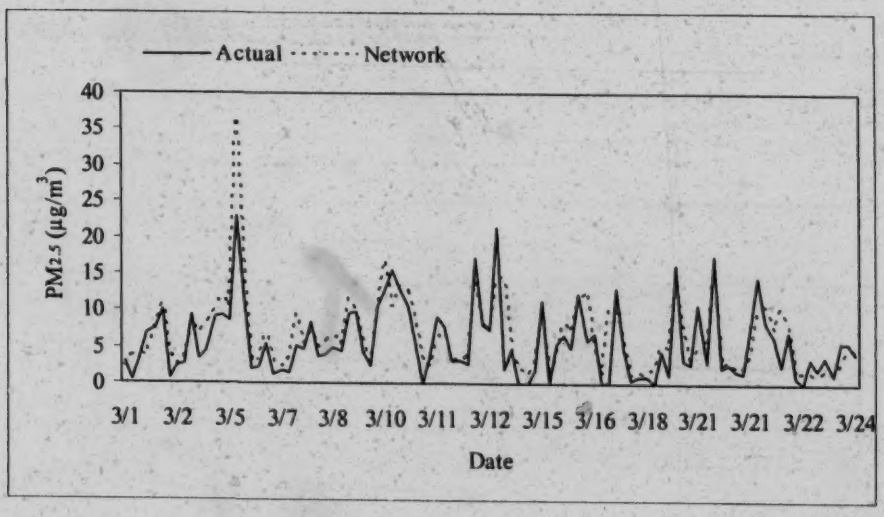


Figure 5- 18 Calgary East spring forecast model with lagged data as an input, $R^2 = 0.63$.

5.3.4.2 Summer model

The summer forecast model developed with lagged data as an input, showed remarkable performance compared to the forecast model developed without lagged data as an input. The model had a R^2 value of 0.78 when developed with 20 hidden layer neurons and 500 training epochs. However, the trend analysis (Figure 5-19) shows that the model had difficulty in following the historical trends. Again, the model appears to be stable which is evident from the Figures B-27 and B-28 shown in Appendix-C. The performance of the forecast model is slightly better than the prediction model when lagged data was included as an input.

5.3.4.3 Fall model

The fall model's performance was similar to that of the spring model with a R^2 value of 0.78.Contrary to the spring model, the fall model was able to follow the higher and lower historical trends with precision. The model was developed with 10 hidden layer neurons and 500 training epochs. However, the prediction had better performance than the forecast model when developed with lagged data as an input. The model appears to be robust which is evident from Figures B-29 and B-30.

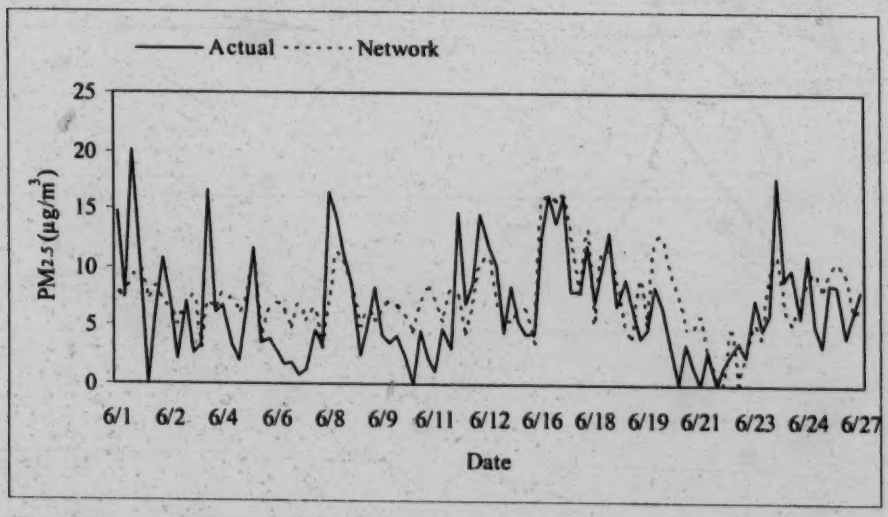


Figure 5- 19 Calgary East summer forecast model with lagged data as an input, R^2 = 0.78

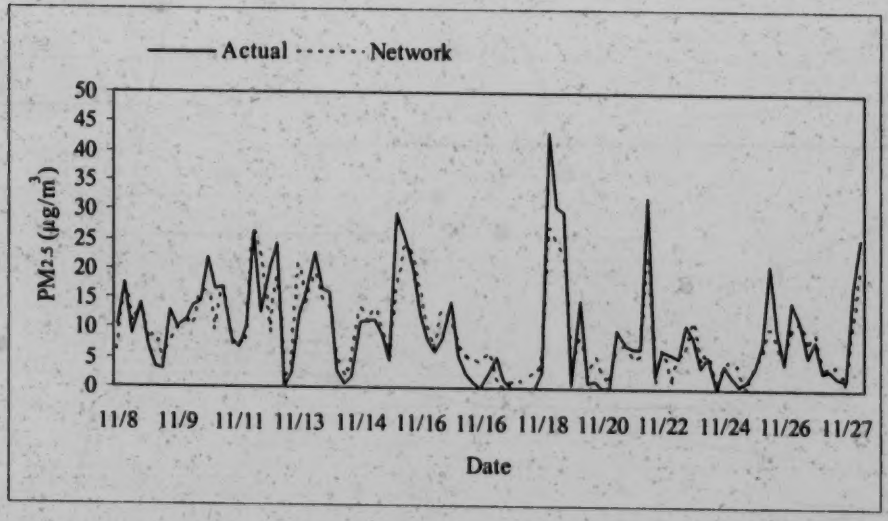


Figure 5- 20 Calgary East fall forecast model with lagged data as an input, $R^2 = 0.78$.

5.3.4.4 Winter model

The winter season model had the worst performance compared to all other models with an R^2 value of 0.37. The winter season model had difficulty in following the historical trends (Figure 5-21). Figure B-31 in Appendix-C shows that the markers were unevenly distributed around the solid line which is indication on unstable model. Also, residual

analyses (Figure B-32) showed some sort of uneven distribution of residuals. Table 5-5 summarizes the forecast model with and without lagged data as an input.

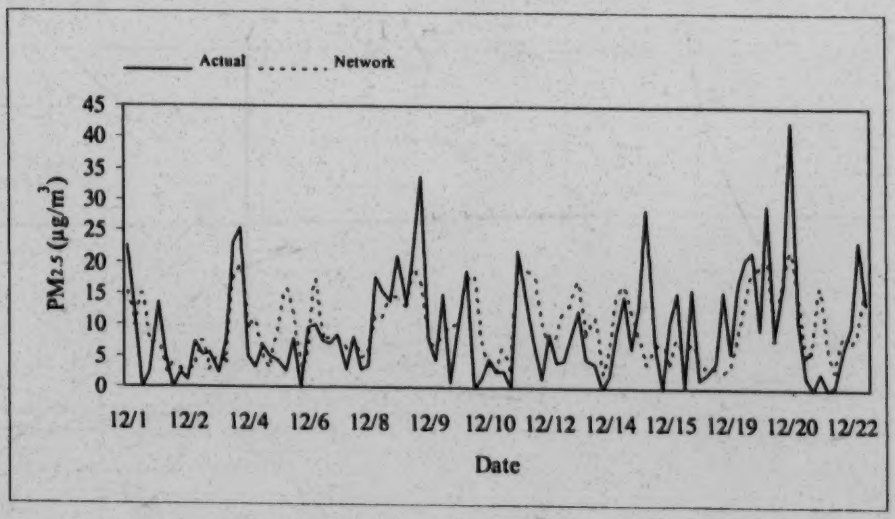


Figure 5- 21 Calgary East winter forecast model with lagged data as an input $R^2 = 0.37$.

Table 5- 5 Summary of forecast model characteristics with and without lagged data as an input.

Architecture	Fall	Winter	Spring	Summer
Forecast model				
No. of hidden layer neurons	10	20	5	5
No. of training epochs	500	200	200	300
Coefficient of multiple determination (R^2)	0.71	0.38	0.67	0.68
Forecast model with lagged data				
No. of hidden layer neurons	10	15	5	20
No. of training epochs	500	500	500	500
Coefficient of multiple determination (R^2)	0.78	0.37	0.63	0.78

5.3.5 *Important Input variables*

Based on the studies done before on $PM_{2.5}$, pollutant parameters such as CO (Perez et al., 2004; Bogo et al., 2003), NO, NO_2 (Harrison et al., 1997), O_3 (McKendry, 2002), SO_2 (Cheng et al., 1998), THC (Harrison et al., 1997) and indexed temporal variables (Su, 2004; McCullum et al., 2004) emerged as core parameters. Keeping the core variables as constants, the meteorological variables were removed one at a time and two at a time and the model is again retrained. The whole input identification exercise was carried out using the optimized prediction model without lagged data for all the seasons.

5.3.5.1 *Spring season*

Opacity (OPA) and Relative Humidity (RH) emerged as an important variable for spring season.

R^2 value decreased from 0.61 to 0.54 and 0.46 for OPA and RH respectively. Photochemical associated particulate formation may be the reason for OPA importance during the spring season. On the other hand, RH is very high in the colder season's which acts as a sink in removing the particles from the atmosphere. Wind direction (WDR) emerged as a predictor variable for $PM_{2.5}$, R^2 value decreased from0.61 to 0.57. Also, there was an inverse correlation between wind speed (WSP) and $PM_{2.5}$ was observed during this season (Table blah). This may be due to localized source emission dominating during this season (Chaloulakou et al., 2003). Meanwhile, WSP, temperature (TEMP) and mixing height (MH), was found to play lesser role in the prediction process. Figure 5-22 shows the change in R^2 value with the removal of input variable.

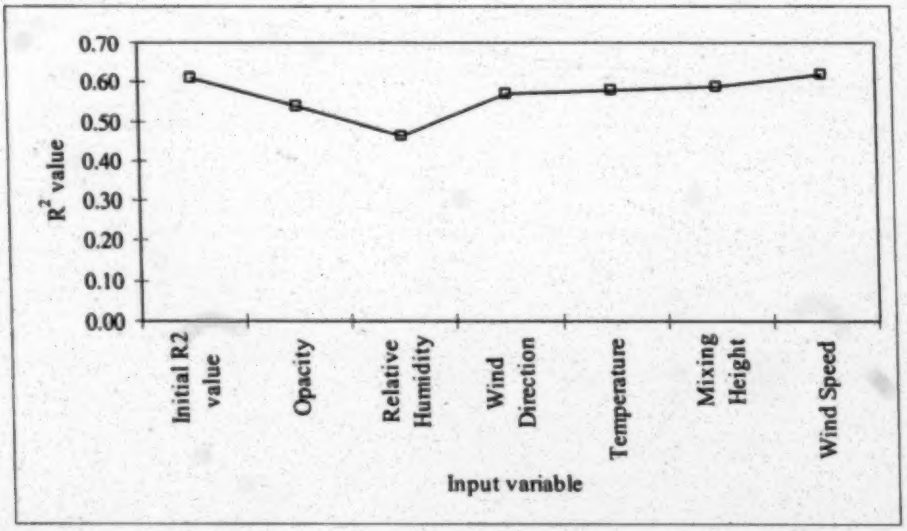


Figure 5- 22 Change in R^2 value due to removal of input variable for spring season model

5.3.5.2 Summer season

OPA, RH and WSP emerged as a major predictor of $PM_{2.5}$ for summer season. Summer season had the maximum $PM_{2.5}$ concentration (14 $\mu g/m^3$) compared to all other seasons. Sunlight playing major role in particulate formation may be the reason for importance of OPA during summer season. WSP, 7.6km/hr was higher compared to other seasons. There were cases of long range of particulates due to high WSP conditions (Fang et al., 1999; McKendry et al., 1999). WSP plays a major role in carrying pollutants during episodic events like forest fires Cheng et al., 1998). RH acts as a scavenger in removing atmospheric particles, and also may act as a surrogate of precipitation events (thundershowers), which is common in summer season.

Figure 5-23 shows the change in R^2 value when a particular variable is removed.

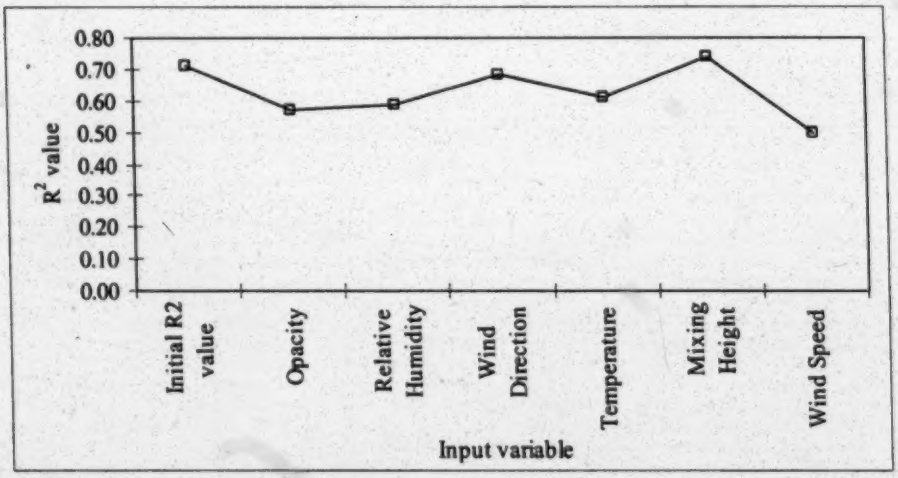


Figure 5- 23 Change in R^2 value due to removal of input variable for summer season model

5.3.5.3 Fall season

The initial R^2 value for fall season was found to be 0.72. RH, OPA, WSP and WDR emerged as key variable for the prediction of $PM_{2.5}$. R^2 value decreased from 0.72 to 0.67 when RH was removed as an input variable. Fall season may have wide suite of sources, which is evident by the emergence of OPA, WSP and WDR as main input variable. Figure 5-24, shows the variation in R^2 based on the removal of particular variable.

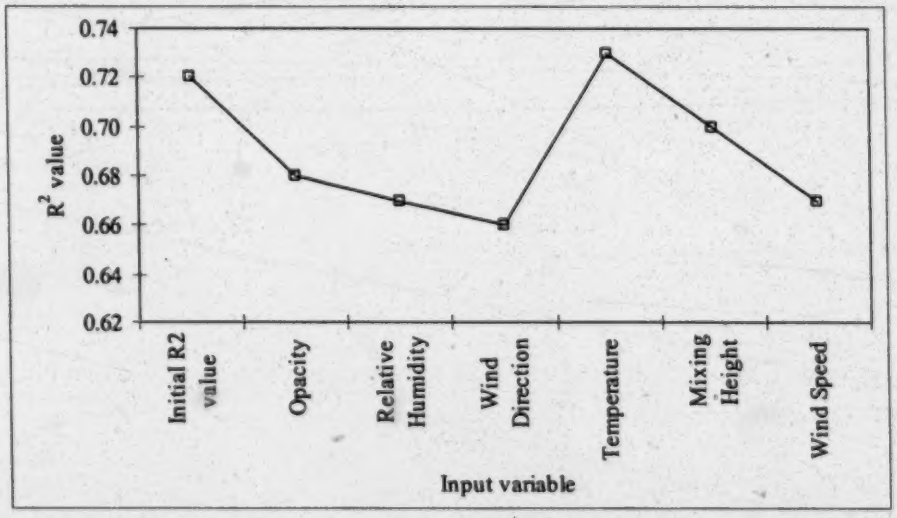


Figure 5- 24 Change in R^2 value due to removal of input variable for fall season model

5.3.5.4 Winter season

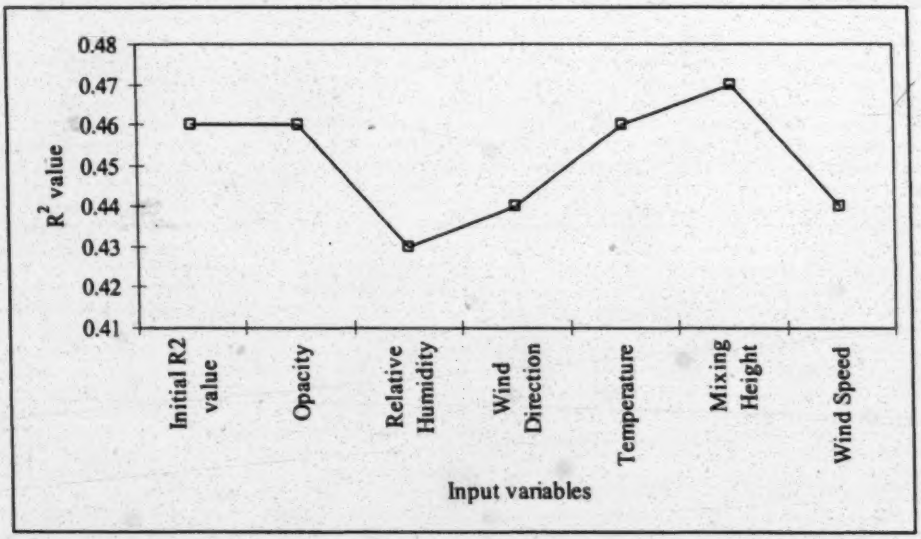


Figure 5- 25 Change in R^2 value due to removal of input variable for winter season model

The winter season had poor prediction ability (R^2 0.46). This reflected on the identification of key input variables. Most of the variables found to play a minimal role in prediction of $PM_{2.5}$. Table 5-6 shows the R^2 value change from the original on removing a particular variable.

Table 5- 6 Change in R^2 value with respect to removal of particular variable

	Spring	**Summer**	**Fall**	**Winter**
Initial R^2 value	0.61	0.71	0.72	0.46
OPA	0.54	0.57	0.68	0.46
RH	0.46	0.59	0.67	0.43
WDR	0.57	0.68	0.66	0.44
MH	0.59	0.74	0.70	0.47
WSP	0.62	0.50	0.67	0.44
TEMP	0.58	0.61	0.73	0.46

Overall, mixing height's contribution towards the prediction was turned out to be negligible for all the seasonal models. The mixing height data used in the present study was calculated using a previous theoretical study done by Benkley and Schulman (1979). Due to the complex terrain conditions, the data calculated may not be representative of the weather conditions prevailing in that area. Meanwhile, temperature emerged an important variable for the summer, and winter seasons. Reason for temperature's unimportance in fall and spring season is not known.

5.4 Conclusions

Two types of season specific models: Prediction and forecast were built using ANNs for the pollutant $PM_{2.5}$ for the city of Calgary. The fall and summer season prediction models performed much better than the spring and winter seasons models. Addition of previous hour $PM_{2.5}$ concentration as an input improved the performance of the prediction models. The developed prediction models can be used as a virtual monitor in case of failures to the real monitors. Also, the models can be used as a tool to check the instrument's performance.

The forecast models which are used to predict the future pollutant concentrations had somewhat lesser performance compared to the prediction models. However, addition of lagged input data did not improve their generalization ability. Relative humidity and opacity emerged as important variables for all seasons in addition to the core variables such as CO, NO, NO_2, O_3, SO_2 and indexed temporal variables.

Future studies are required to address the poor prediction ability of the winter season models. Using meteorological pre-processors such as CALMET (Senes consultants Limited, 1996) which incorporates the terrain features may address the problem of representative mixing height data. Simulation of Chinook wind conditions as an input into the model may improve the winter season's prediction ability. More studies are required by including previous days $PM_{2.5}$ and elemental ion compounds as an input to improve the generalization ability of the models.

5.5 References

Abdul-Wahab, S. A., and Al-Alawi, S.M., 2002: Assessment and prediction of tropospheric ozone concentration levels using Artificial Neural Network. *Environmental Modelling and Software*, **17,** 219-228.

Angle, R. P., and Sakiyama, S.K., 1991: *Plume dispersion in Alberta*. Standards and Approvals Division, Alberta Environment, 428 pp. (ISBN 0-7732-1062-8)

Ashbaugh, L. L., 1983: A statistical trajectory technique for determining air pollution source regions. *J. Air Pollut. Control Assoc.*, **33,** 1096-1098.

Benkley, C. W., and Schulman, L. L., 1979: Estimating hourly mixing depths from historical meteorological data. *J. of App. Meteor.*, **18,** 772-780.

Bogo, H., Otero, M., Castro, P., Ozafran, M.J., Kreiner, A., Calvo, E.J., and Negri, R. M., 2003: Study of atmospheric particulate matter in Buenos Aires City. *Atmos. Environ.*, **37,** 1135-1147.

Chaloulakou, A., Grivas, G., and Spyrellis, N., 2003a: Neural Network and multiple regression models for PM_{10} prediction in Athens: A comparative assessment. *J. Air & Waste Management Association*, **53,** 1183-1190.

Chaloulakou, A., Kassomenos, P., Spyrellis, N., Demokritou, P., and Koutrakis, P., 2003b: Measurement of PM_{10} and $PM_{2.5}$ particle concentrations in Athens, Greece. *Atmos. Environ.*, **37,** 649-660.

Chelani, A. B., Rajghate, D.G., and Hasan, M.Z., 2002: Prediction of ambient PM_{10} and toxic metals using Artificial Neural Networks. *J. Air & Waste Management Association*, **52,** 805-810.

Cheng, L., Sandhu, H.S., Angle, R.P., and Myrick, R.H., 1998: Characteristic of inhalable particulate matter in Alberta Cities. *Atmos. Environ.*, **32,** 3835-3844.

Cheng, L.; McDonald, K.M.; Angle, R.P.; and Sandhu, H.S., 1998. Forest Fire Enhanced Photochemical Air Pollution- A Case Study; *Atmos. Environ.* **32**, 673-681.

El-Din, A. G., and Smith, D.W., 2002: A neural network model to predict the wastewater inflow incorporating rainfall events. *Water Res.*, **36,** 1115-1126.

Environment Canada, 2000: Canadian Environmental Protection Act, 1999, priority substances list assessment report- respirable particulate matter less than or equal to 10 Microns. Cat. No. En40-215/47E, 73 pp. (Available from www.ec.gc.ca)

Fang, M., Zheng, M., Wang, F., Chim, K.S., and Kot, S.C., 1999. The Long-range Transport of Aerosols from Northern China to Hong Kong—a multi technique study. *Atmos. Environ.* **33**, 1803–1817.

Fernandez de Castro, B. M., Sanchez, J.M.P., Manteiga, W.G., Bande, M.F., Bermudez Cela, J.R., and Fernandez, J.J.H, 2003: Prediction of SO_2 levels using Neural Networks. *J. Air & Waste Managment Association*, **53,** 532-539.

Gardner, M. W., and Dorling, S.R., 1998: Artificial Neural Networks (The Multilayer Perceptron) - A review of applications in the atmospheric sciences. *Atmos. Environ.*, **32,** 2627-2636.

Gardner, M.W., and Dorling, S.R., 1999: Modeling and prediction of hourly NO_x and NO_2 concentrations in urban air in London. *Atmos. Environ.*, **33,** 709-719.

Gardner, M.W., and Dorling, S.R., 2001: Artificial Neural Network-derived trends in daily maximum surface ozone concentrations. *J. Air & Waste Managment Association*, **51,** 1202-1210.

Harrison, R. M., Deacon, A.R., and Jones, M.R., 1997: Sources and processes affecting concentrations of PM_{10} and $PM_{2.5}$ particulate matter in Birmingham (U.K.). *Atmos. Environ.*, **31,** 4103-4117.

Hasham, F. A., Kindzierski, W.B., and Stanley, S.J., 2004: Modeling of hourly NO_x concentrations using Artificial Neural Networks. *J. Environ. Eng. Sci.*, **3,** 111-119.

Hies, T., Treffeisen, R., Sebald, L., and Reimer, E, 2000: Spectral analysis of air pollutants, Part 1: elemental carbon time series. *Atmos. Environ.*, **34,** 3495-3502.

Hopke, P.K., Gladney, E.S., Gorden, G.E., Zoller, W.H., and Jones, A.G., 1976: The use of multivariate analysis to identify sources of selected elements in Boston urban aerosol. *Atmos. Environ.*, **10,** 1015-1025.

McCullum, K., Kindzierski, W.B., Gamal El-Din, M., and Myrick, B., 2004: Elemental trends in urban ambient particulate matter in downtown Edmonton and Calgary, Alberta. *Cold Region Engineering Conference proceedings C.D.*, Edmonton, Canada (Available from the Construction Research Institute of Canada, #310G Civil Electrical Engineering Building, University of Alberta, Edmonton, Alberta, Canada, T6G 2G7)

McKendry, I. G., 2002: Evaluation of Artificial Neural Networks for fine particulate pollution (PM_{10} and $PM_{2.5}$) forecasting. *J. Air & Waste Managment Association*, **52,** 1096-1101.

McKendry, I. G., Hacker, J. P., Stull, R., Sakiyama, S., Mignacca, D., and Reid, K., 2001. Long range transport of Asian Dust to the Lower Fraser Valley, British Columbia, Canada. *J. of Geophys. Res.-atm.*, **106,** 18361-18370.

Nkemdirim, 1988: An assessment of the relationship between functional groups of weather elements and atmospheric pollution in Calgary, Canada. *Atmos. Environ.*, **22,** 2287-2296.

NRC, 1998: *Research priorities for airborne particulate matter*. Vol. 1, National Academic Press, 195 pp.

Perez, P., Palacios, R., and Castillo, A., 2004: Carbon Monoxide concentration forecasting in Santiago, Chile. *J. Air & Waste Managment Association*, **54,** 908-913.

Schwartz, J., Dockery, D.W., and Neas, L.M., 1996: Is daily mortality associated specifically with fine particles? *J. Air & Waste Managment Association*, **46,** 927-939.

Senes Consultants limited, 1996: A mixing height study for North America (1987-1991). Report prepared for the Atmospheric Environment Service, North York, Ontario, 52 pp. (Available from Air quality Research branch, Atmospheric Environment Service, 4905 Dufferin Street, North York, Ontario M3H 5T4)

Su, D., Gamal El-Din, M., and Selvaraj, M., 2004: *Modelling $PM_{2.5}$ in Calgary and Edmonton: Application of Artificial Neural Networks.* Report prepared for the air section and science standards branch, Alberta Environment, Edmonton, Alberta, 61 pp. (Available from Air section and science Standards branch, Alberta Environment, 9820-106 street, Edmonton, Alberta, T6G 2M8.)

Su, D., 2004: Application of Artificial Neural Network for Modelling Ground level Ozone concentration in Calgary and Edmonton. M.Sc thesis, Department of Civil and Environmental Engineering, University of Alberta, 209 pp. (Available from University of Alberta, Edmonton, Canada.)

Sunyer, J., Castellsague, J., Saez, M., Tobias, A., and Anto, J.M., 1996: Air pollution and mortality in Barcelona. *J. Epidemiol. Commun. Health*, **50,** S76-S80.

6.0 FINDINGS AND RECOMMENDATIONS

The following are the findings from this research

- $PM_{2.5}$ exhibit a seasonal variation in Edmonton, subdividing the historical data according to seasons such as Spring (February, March and April), Summer (May, June and July), Fall (August, September and October), and Winter (November, December and January) improved the prediction ability of the models.

- Two types of season specific models were successfully built using neural networks for the cities of Edmonton and Calgary

- The summer season models for the city of Edmonton showed poor performance compared to other season models. The exact reasons for this are unknown.

- The winter season model for the city of Calgary showed poor prediction capability compared to the other season models. This may be due to the erratic winter weather prevalent in Calgary.

- The systematic approach of model modeling followed in this research gave better results than the other models in the literature.

- Addition of 1-hr lagged $PM_{2.5}$ data improved the model's performance, which shows that $PM_{2.5}$ concentration in Edmonton and Calgary is persistent.

Some of the recommendations for future research are:

- Kohonen network should be used to divide the data into seasons according to the meteorological data.

- Simulation of Chinook wind condition as an input to the Calgary model may improve the performance of winter model.

- Addition of precipitation data as an input should be investigated to try and improve the model's performance.

- More detailed study to ascertain the sources of $PM_{2.5}$ using ANNs should be done.

Appendix-A

EDMONTON EAST MODELS STABILITY CHECK

VIRTUAL MONITOR MODELS

Spring season

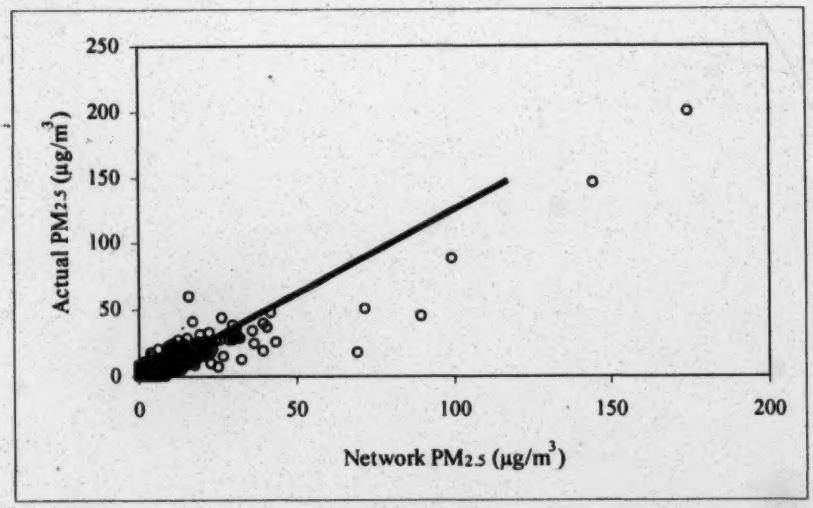


Figure A- 1 Network vs. Actual $PM_{2.5}$ for Edmonton East spring virtual monitor model.

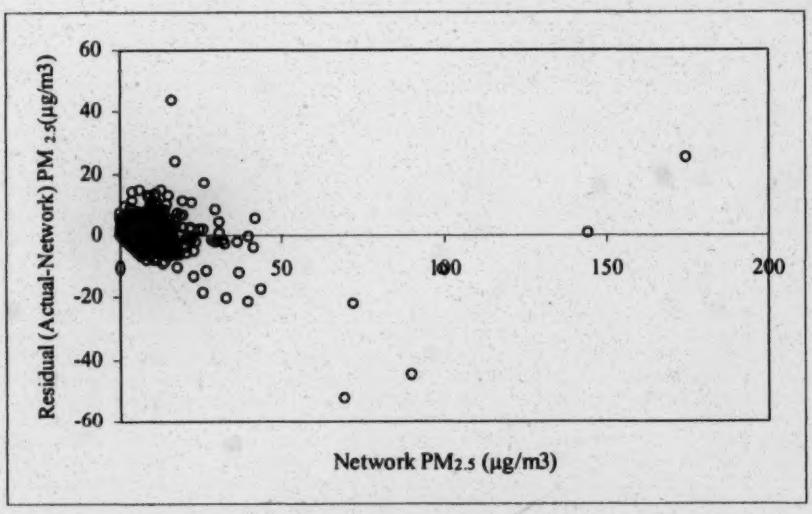


Figure A- 2 Edmonton East spring virtual monitor residual analysis

Summer season

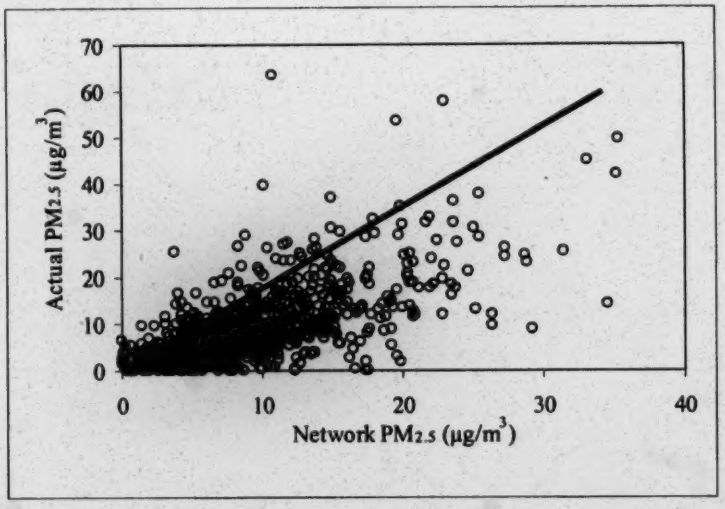


Figure A- 3 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model.

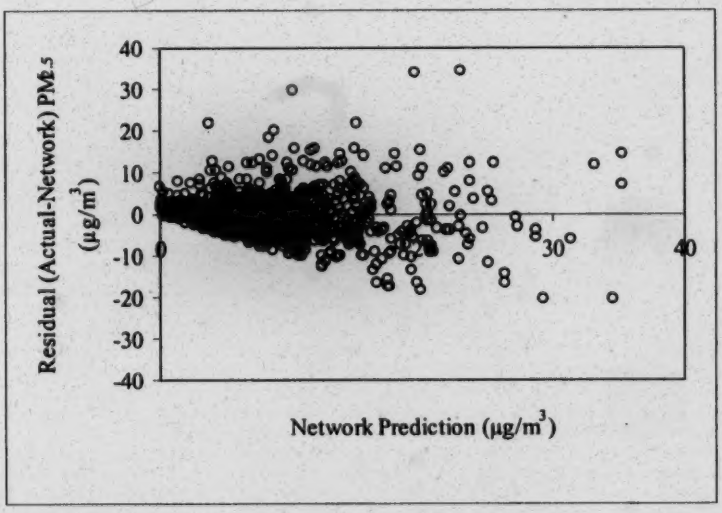


Figure A- 4 Edmonton East summer virtual monitor residual analysis

Fall season

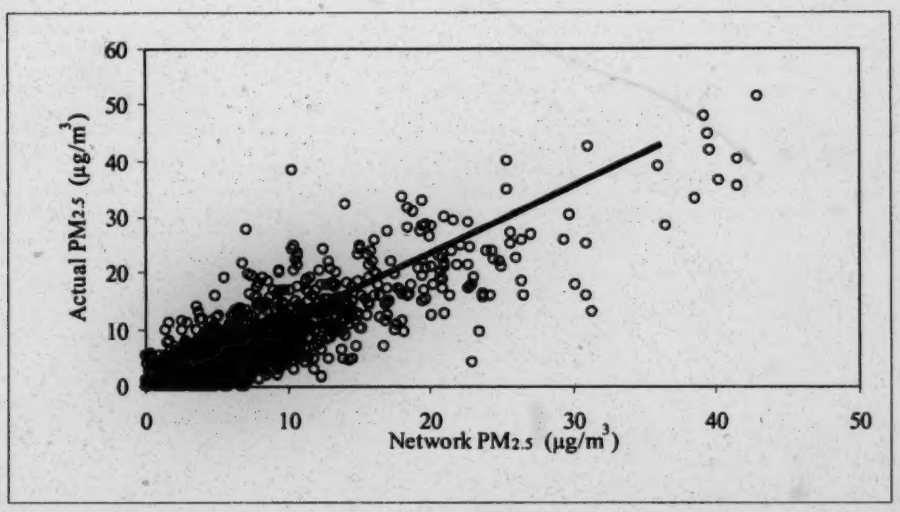


Figure A- 5 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model.

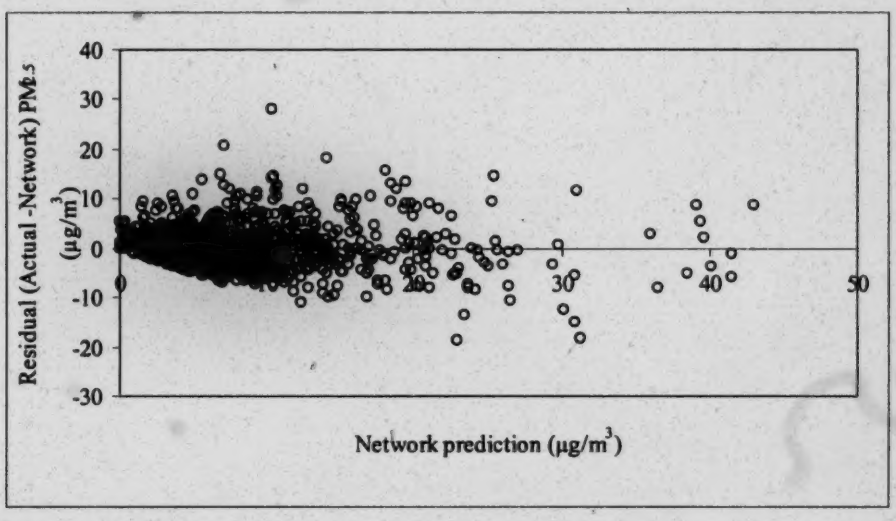


Figure A- 6 Edmonton East fall virtual monitor residual analysis

Winter season

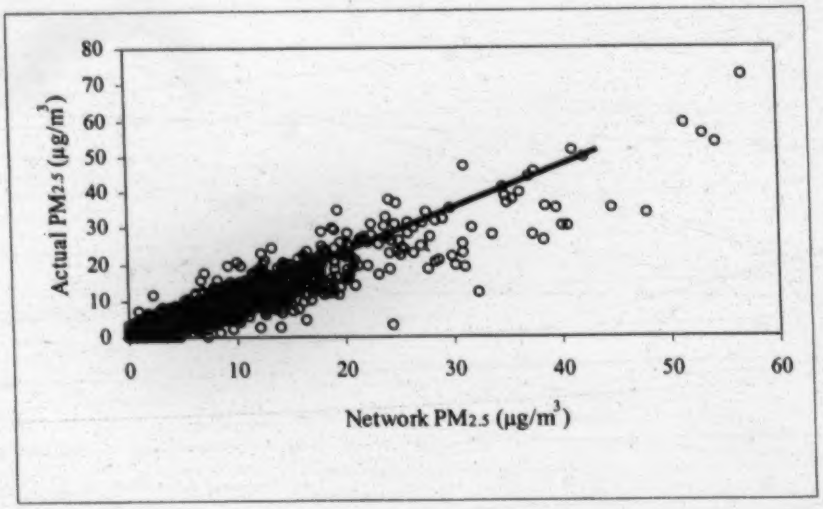


Figure A- 7 Network vs. Actual $PM_{2.5}$ for Edmonton East summer virtual monitor model.

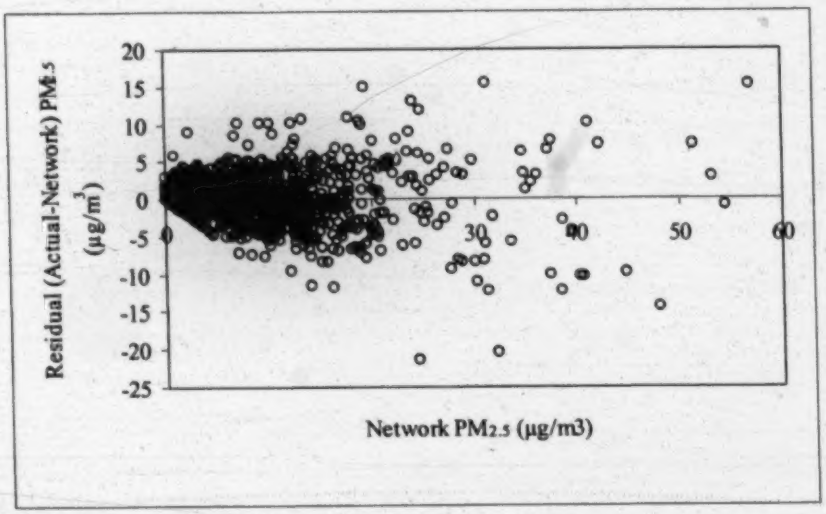


Figure A- 8 Edmonton East winter virtual monitor residual analysis

FORECAST MODELS

Spring season

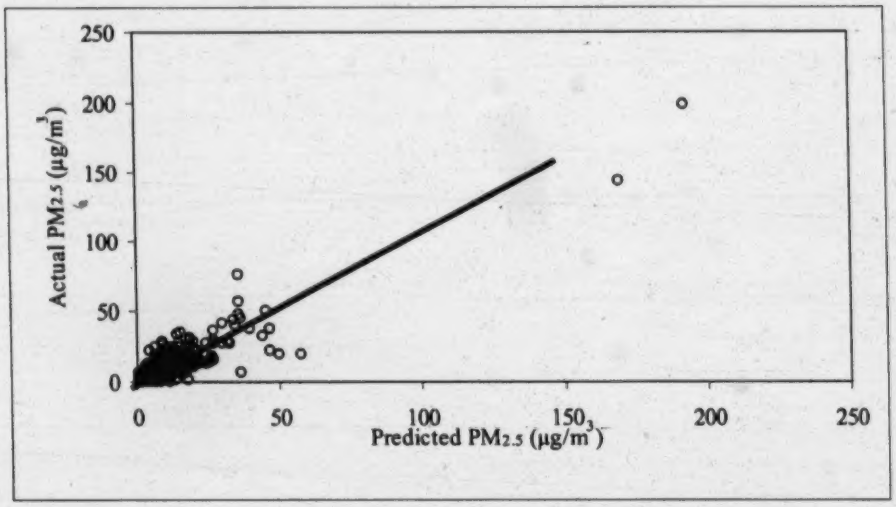


Figure A- 9 Network vs. Actual $PM_{2.5}$ for Edmonton East spring forecast model.

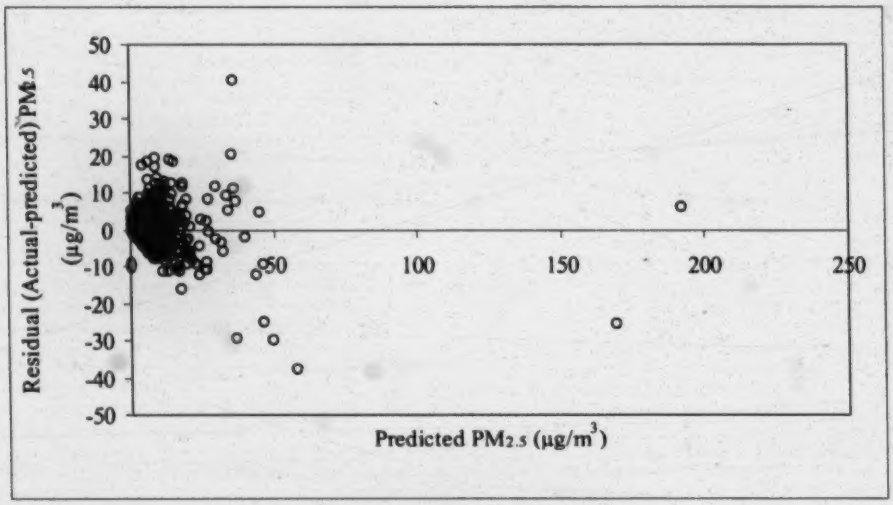


Figure A- 10 Edmonton East spring forecast residual analysis

Summer season

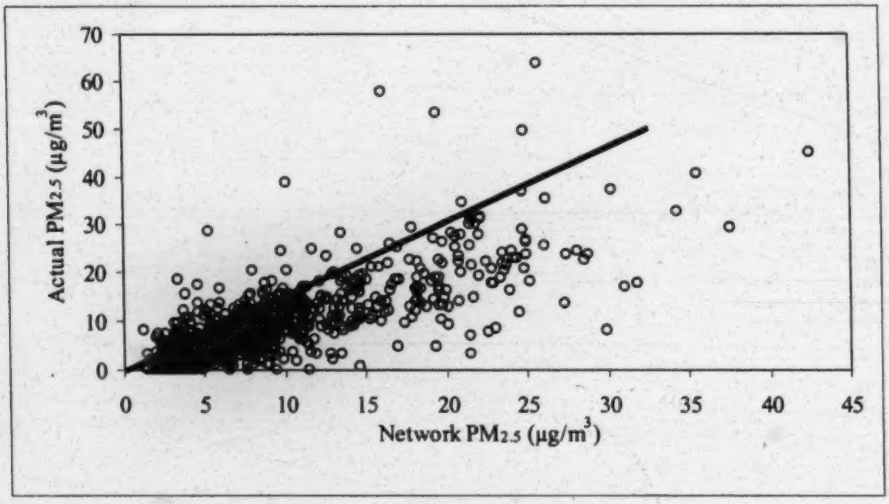


Figure A- 11 Network vs. Actual $PM_{2.5}$ for Edmonton East summer forecast model.

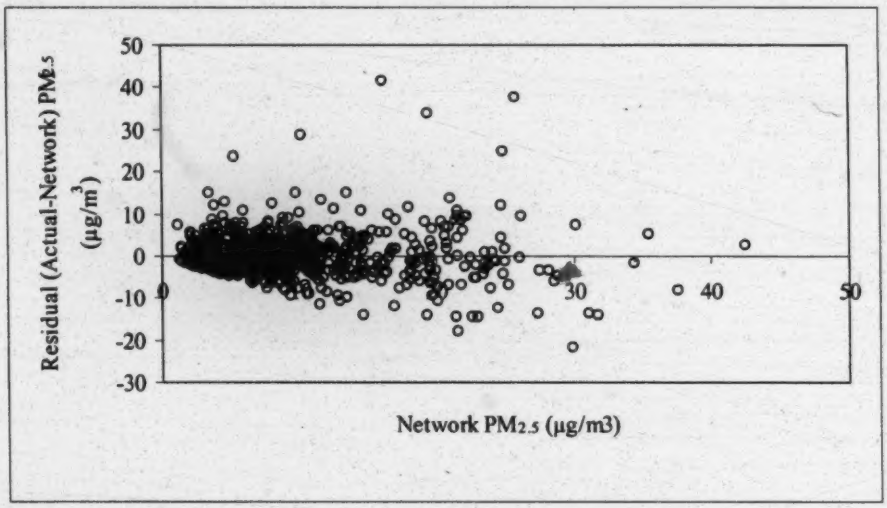


Figure A- 12 Edmonton East summer forecast residual analysis

Fall season

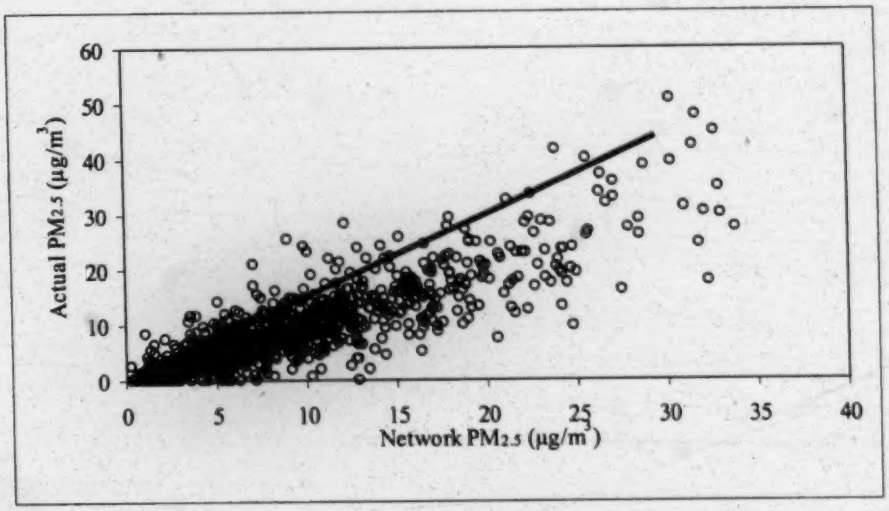


Figure A- 13 Network vs. Actual $PM_{2.5}$ for Edmonton East fall forecast model.

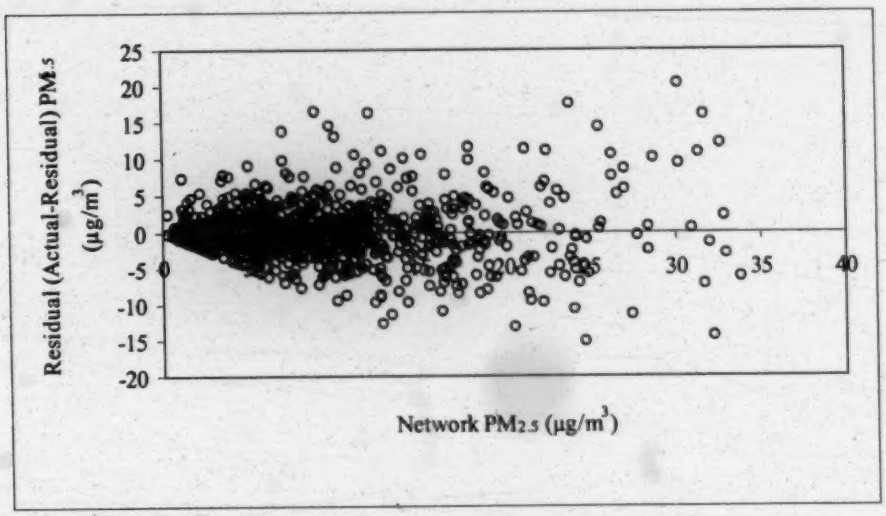


Figure A- 14 Edmonton East fall forecast residual analysis

Winter season

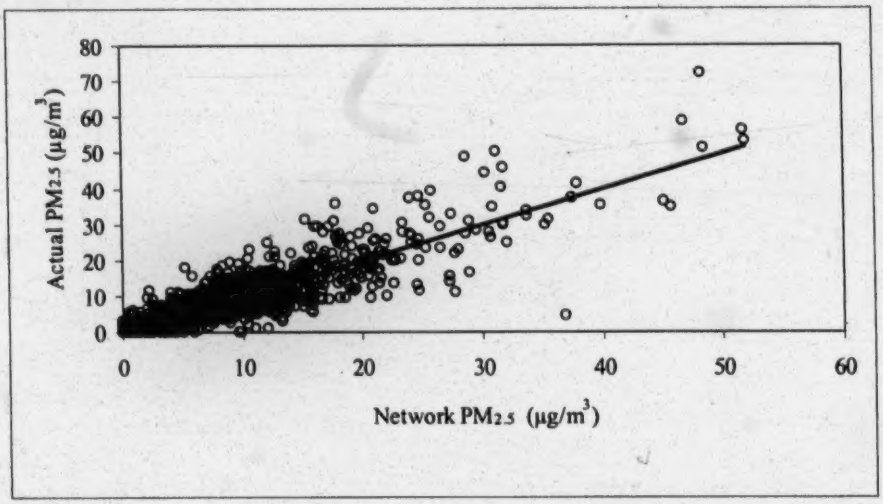


Figure A- 15 Network vs. Actual $PM_{2.5}$ for Edmonton East winter forecast model.

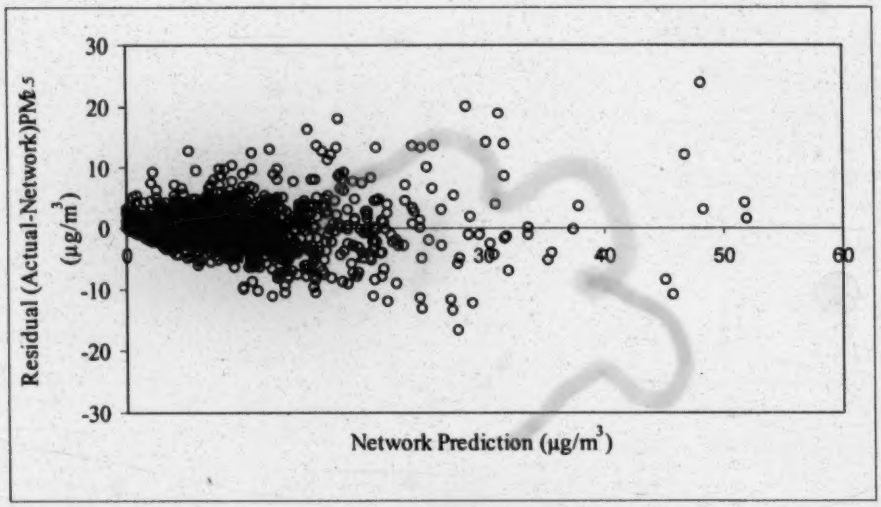


Figure A- 16 Edmonton East winter forecast residual analysis

APPENDIX-B

CALGARY EAST MODELS STABILITY CHECK

PREDICTION MODELS

Spring season

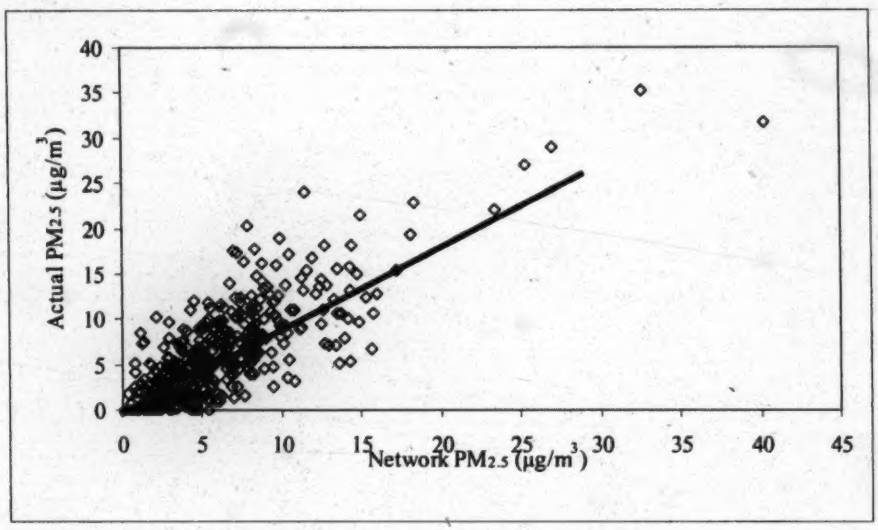


Figure B- 1 Network vs. Actual $PM_{2.5}$ for Calgary spring prediction model.

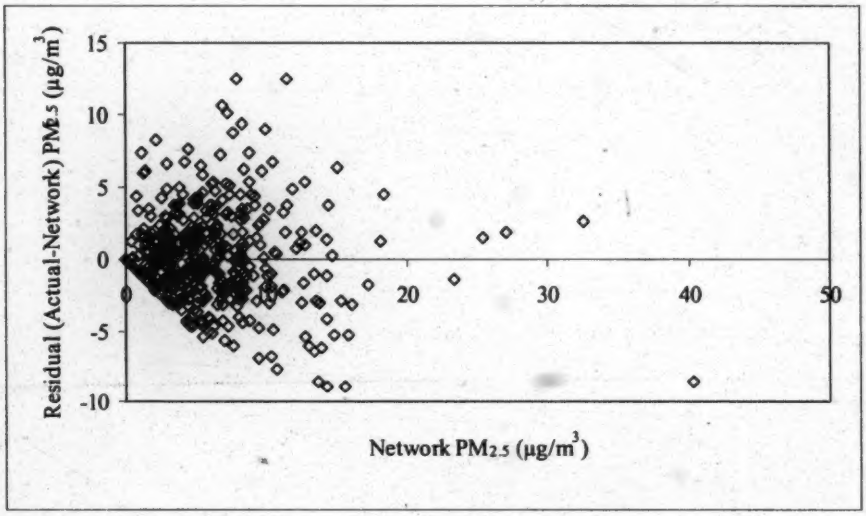


Figure B- 2 Calgary East spring prediction model residual analysis.

Summer season

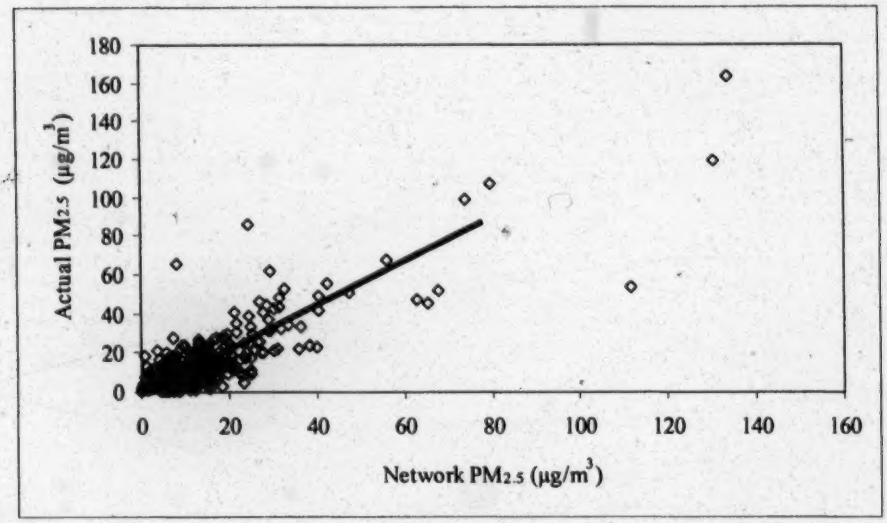


Figure B- 3 Network vs. Actual $PM_{2.5}$ for Calgary summer prediction model

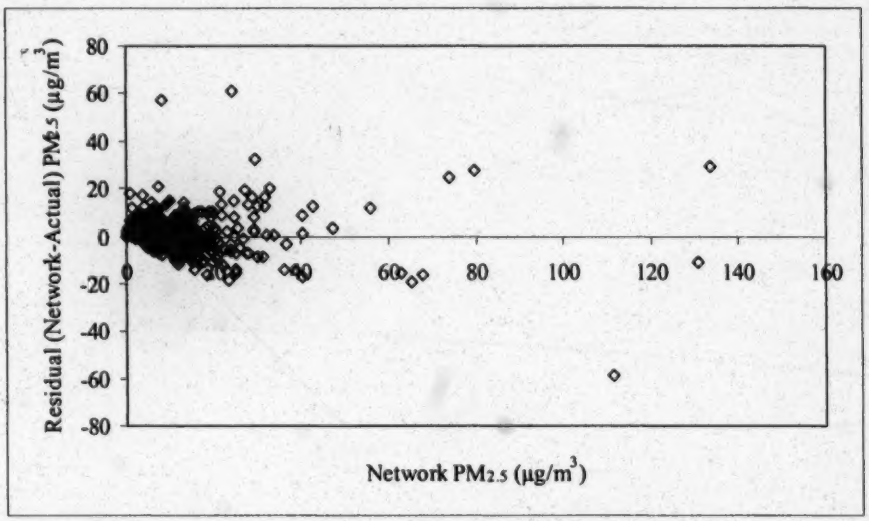


Figure B- 4 Calgary East summer prediction model residual analysis.

Fall season

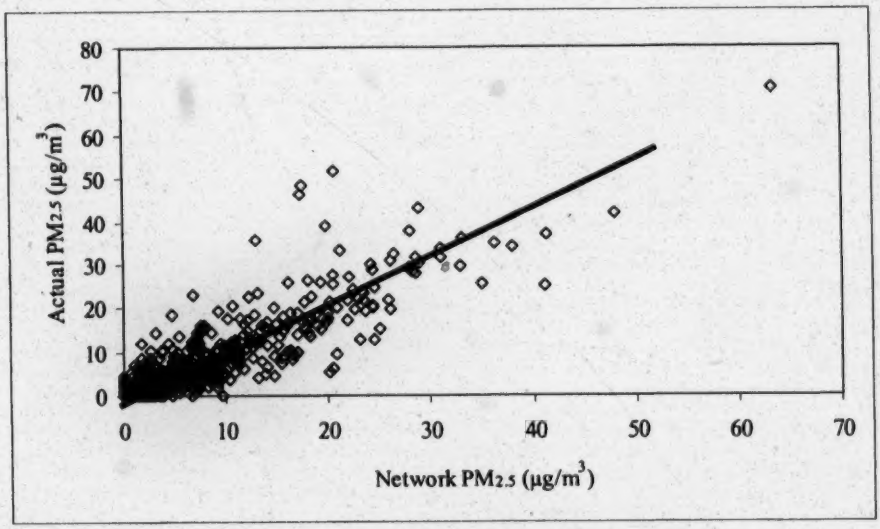


Figure B- 5 Network vs. Actual $PM_{2.5}$ for Calgary fall prediction model

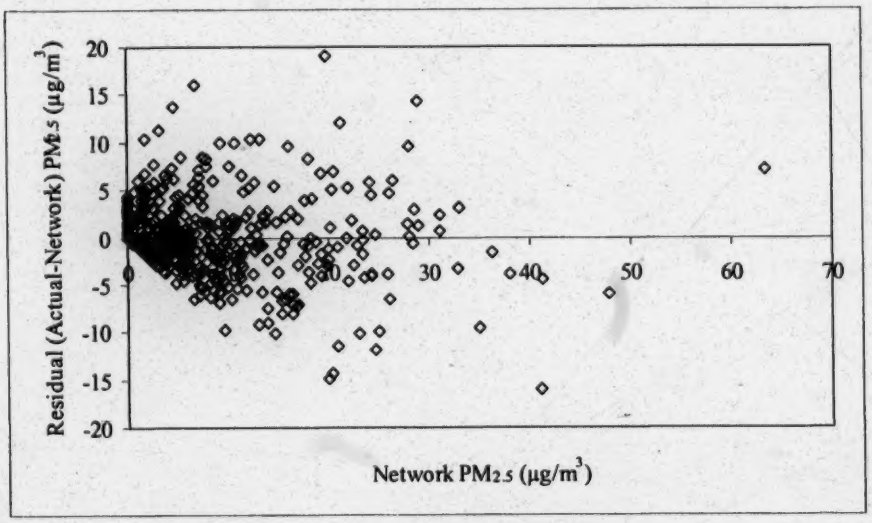


Figure B- 6 Calgary East fall prediction model residual analysis.

Winter season

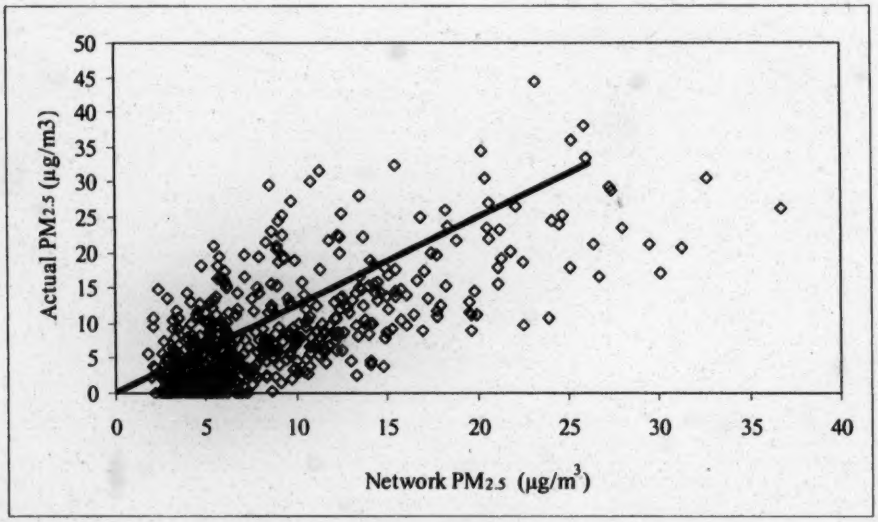


Figure B- 7 Network vs. Actual $PM_{2.5}$ for Calgary winter prediction model.

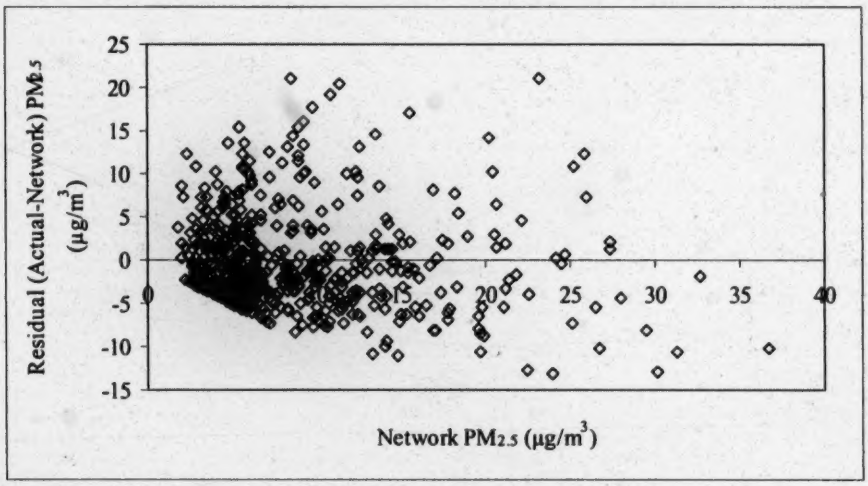


Figure B- 8 Calgary East winter prediction model residual analysis.

PREDICTION MODELS WITH LAGGED DATA AS AN INPUT:

Spring season

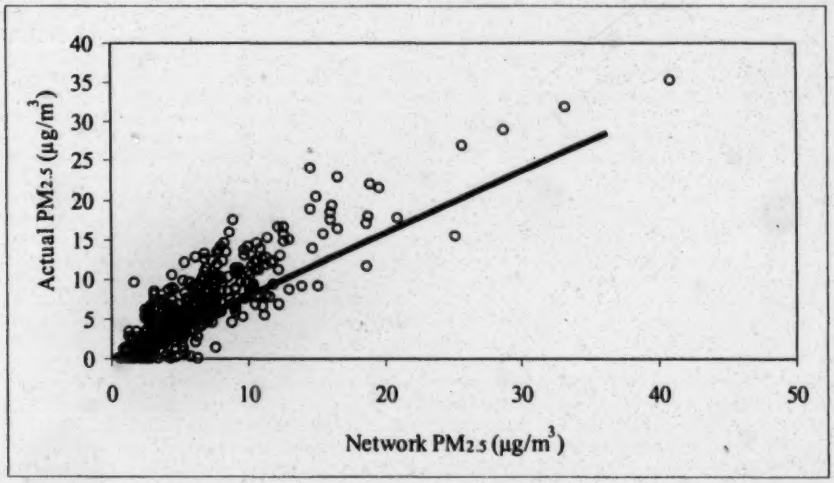


Figure B- 9 Network vs. Actual $PM_{2.5}$ for Calgary spring prediction model with lagged data as an input.

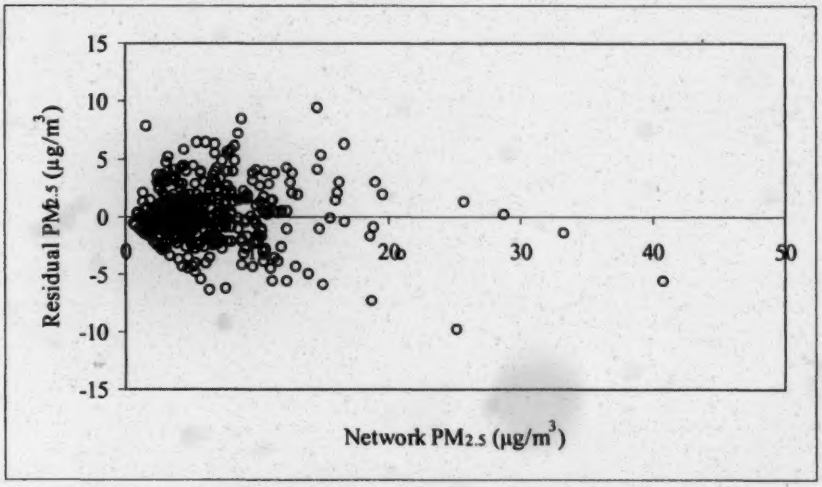


Figure B- 10 Calgary East spring prediction model (lagged data) residual analysis.

Summer season

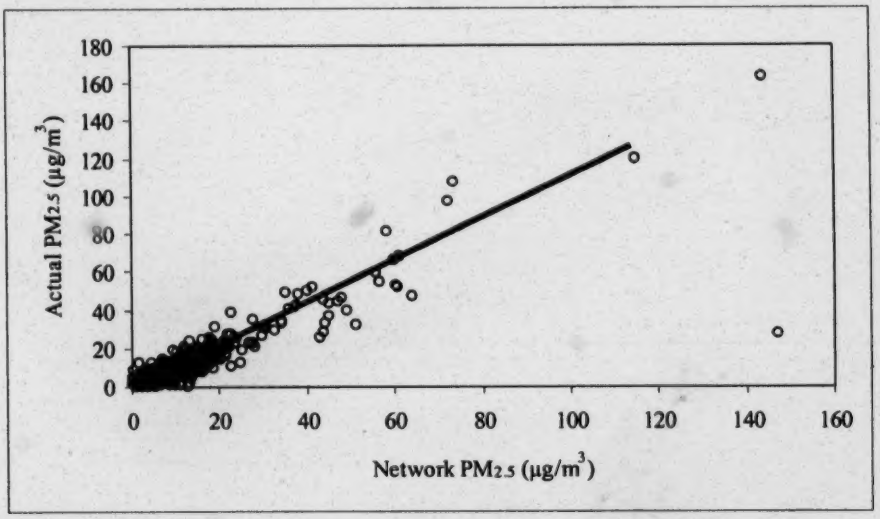


Figure B- 11 Network vs. Actual $PM_{2.5}$ for Calgary summer prediction model with lagged data as an input.

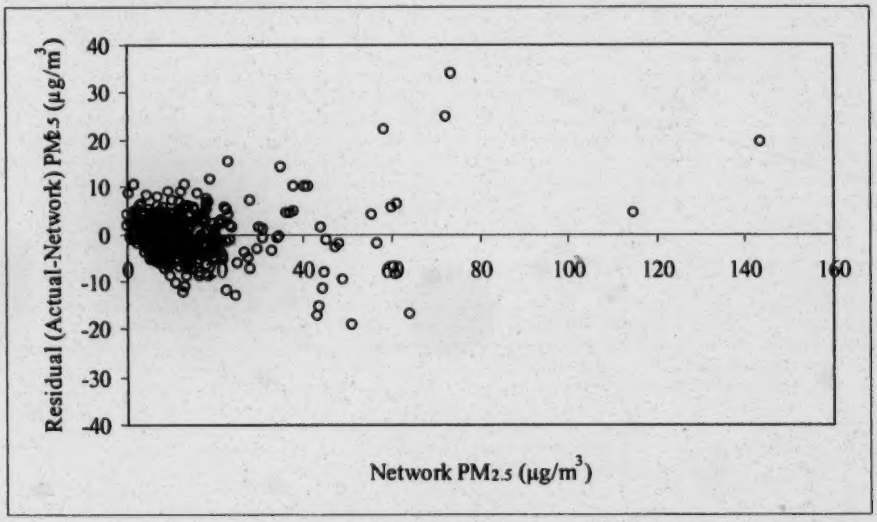


Figure B- 12 Calgary East summer prediction model (lagged data) residual analysis.

Fall season

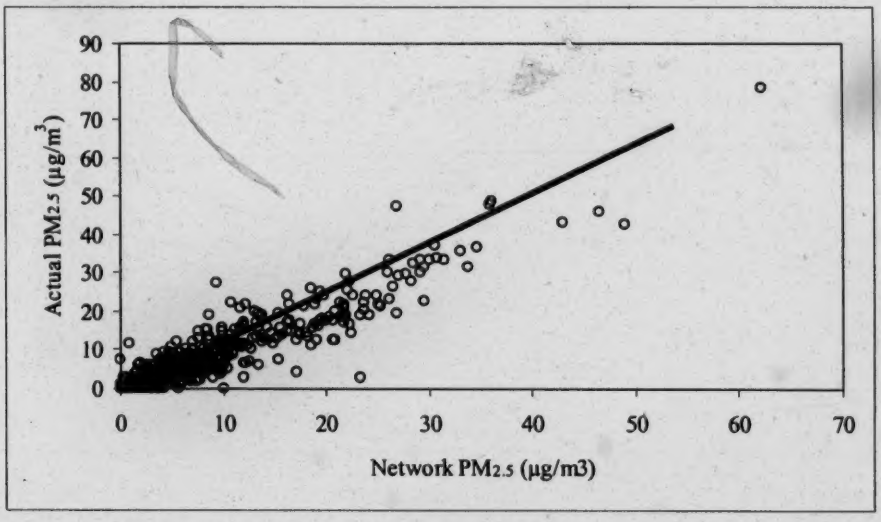


Figure B- 13 Network vs. Actual $PM_{2.5}$ for Calgary fall prediction model with lagged data as an input.

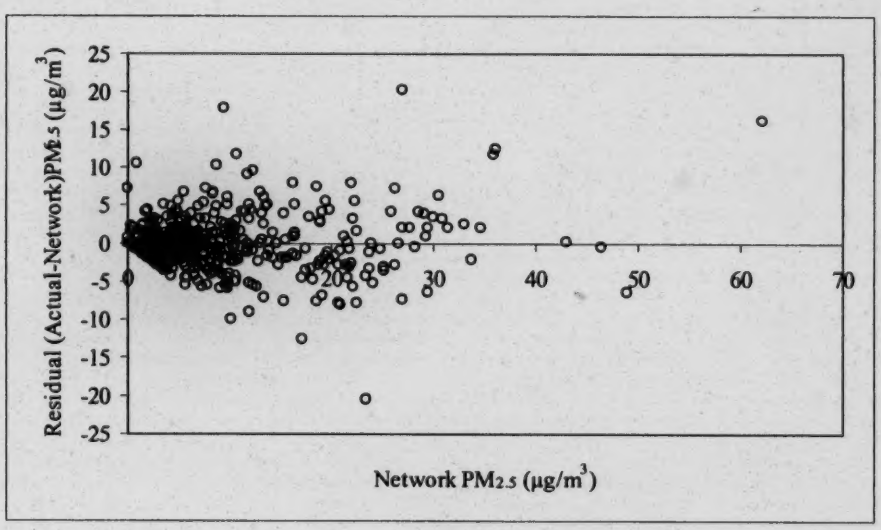


Figure B- 14 Calgary East fall prediction model (lagged data) residual analysis.

Winter season

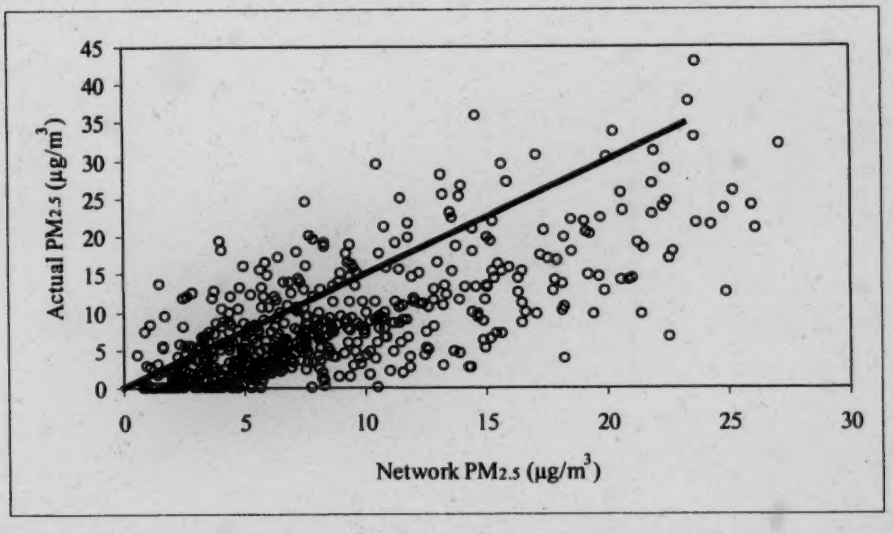


Figure B- 15 Network vs. Actual $PM_{2.5}$ for Calgary winter prediction model with lagged data as an input.

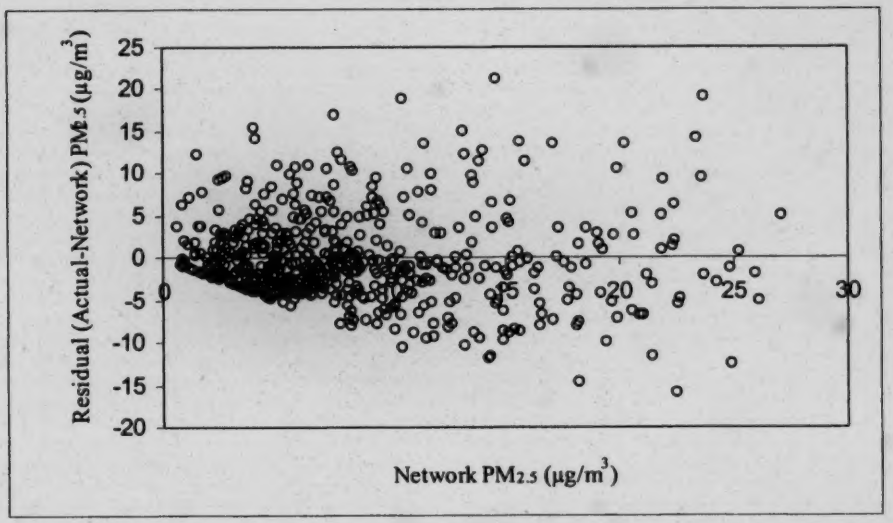


Figure B- 16 Calgary East winter prediction model (lagged data) residual analysis.

FORECAST MODEL

Spring season

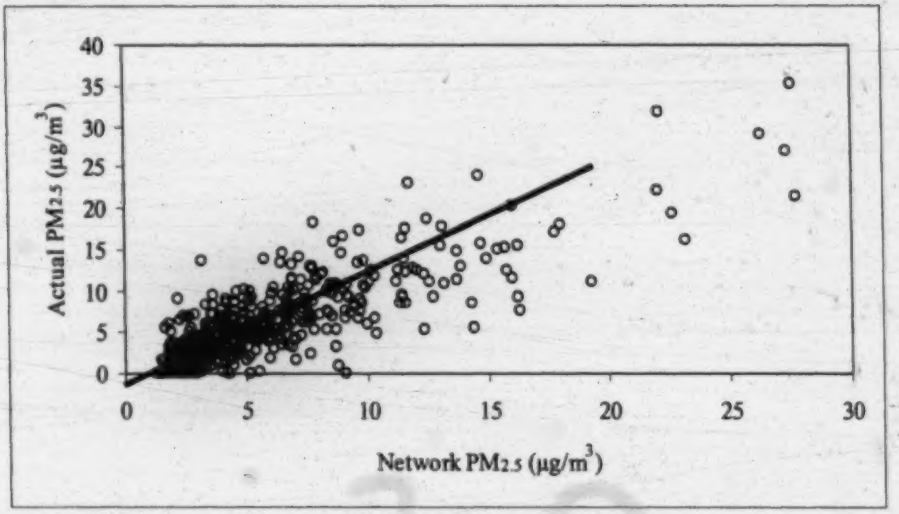


Figure B- 17 Network vs. Actual $PM_{2.5}$ for Calgary spring forecast model.

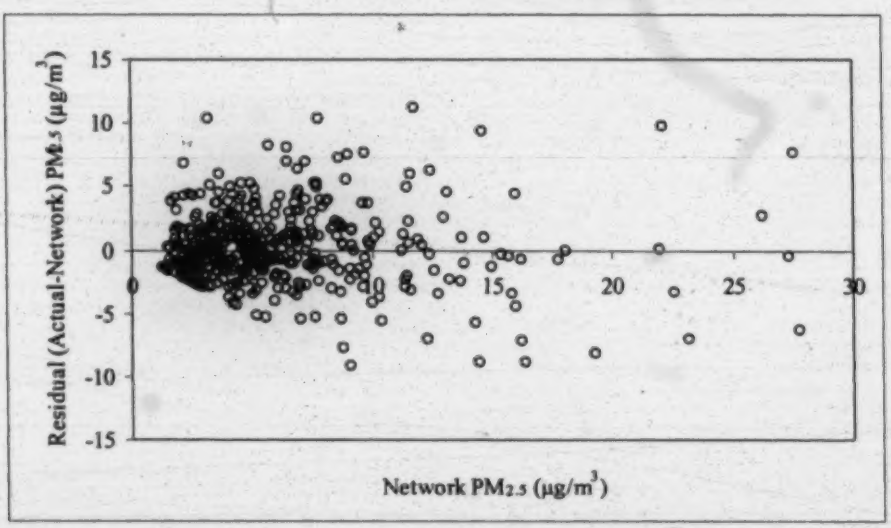


Figure B- 18 Calgary East spring forecast model residual analysis.

Summer season

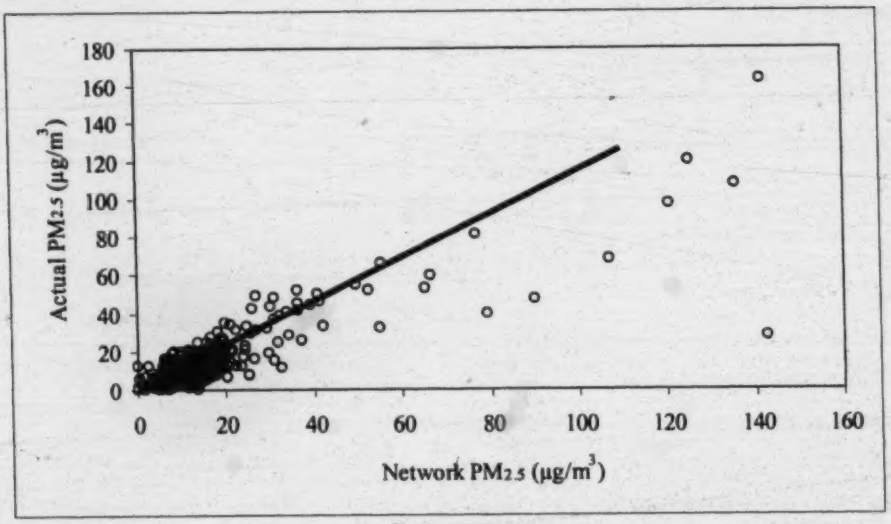


Figure B- 19 Network vs. Actual $PM_{2.5}$ for Calgary East summer forecast model.

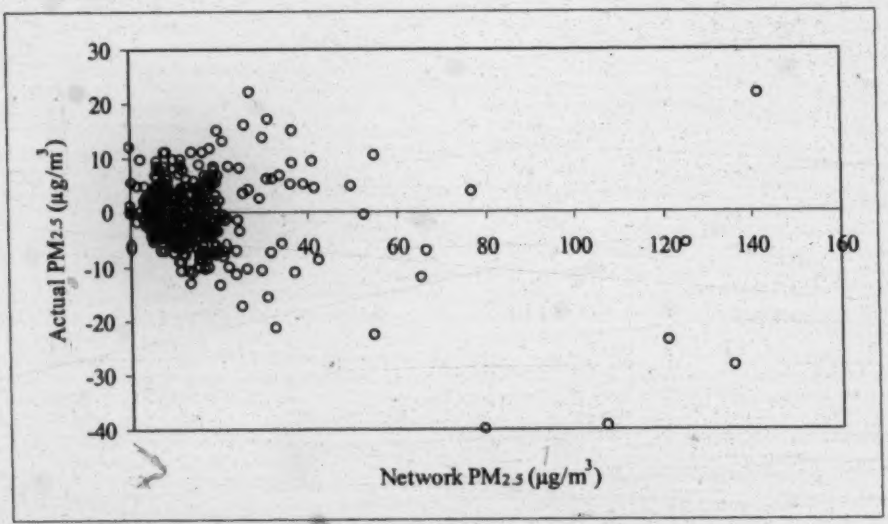


Figure B- 20 Calgary East summer forecast model residual analysis.

Fall season

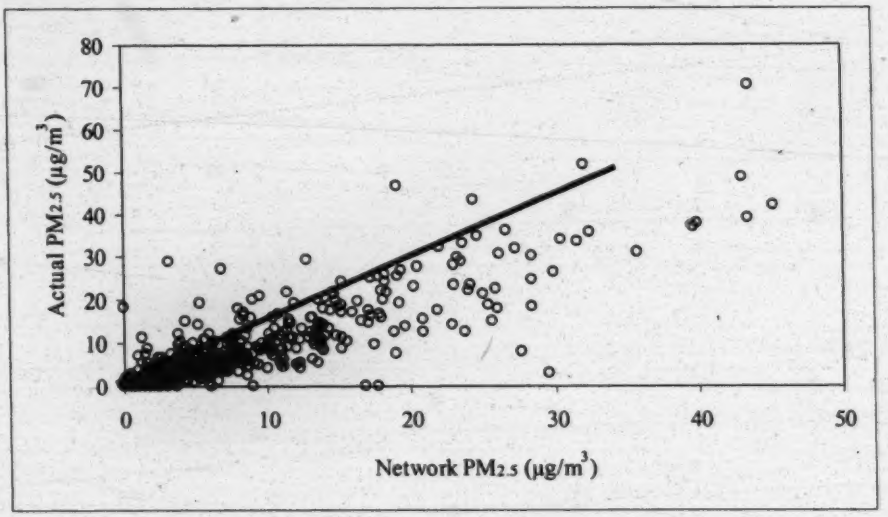


Figure B- 21 Network vs. Actual $PM_{2.5}$ for Calgary East fall forecast model.

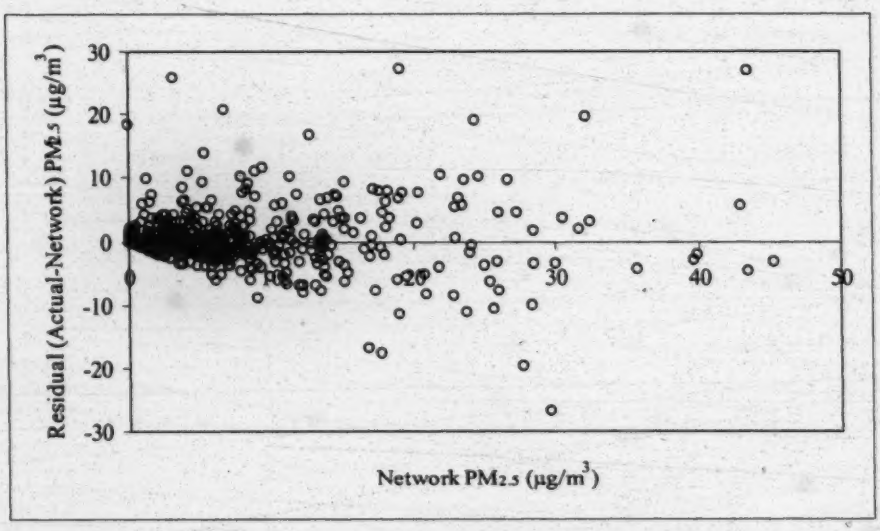


Figure B- 22 Calgary East fall forecast model residual analysis.

Winter season

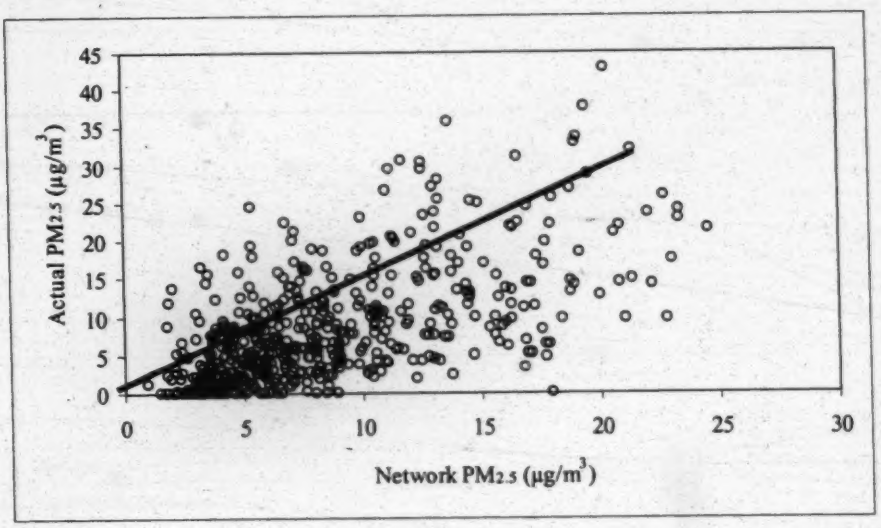


Figure B- 23 Network vs. Actual $PM_{2.5}$ for Calgary East winter forecast model.

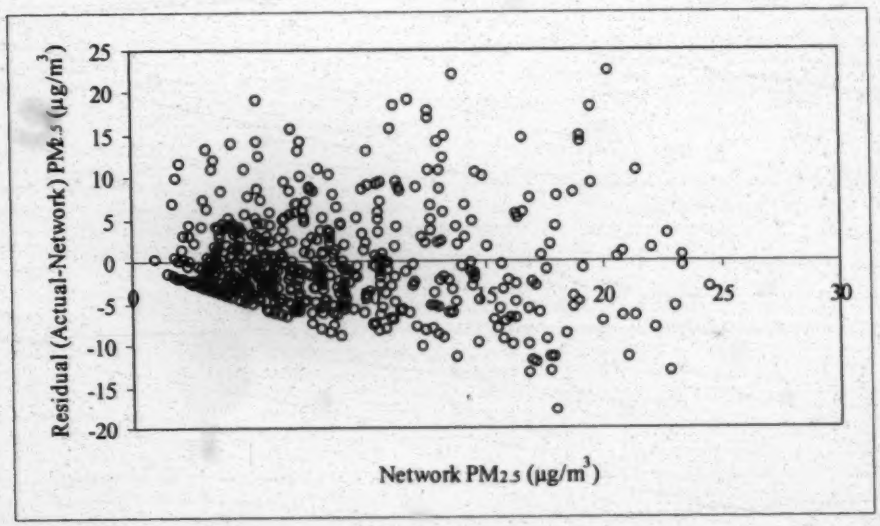


Figure B- 24 Calgary East winter forecast model residual analysis.

FORECAST MODEL WITH LAGGED DATA AS AN INPUT:

Spring season

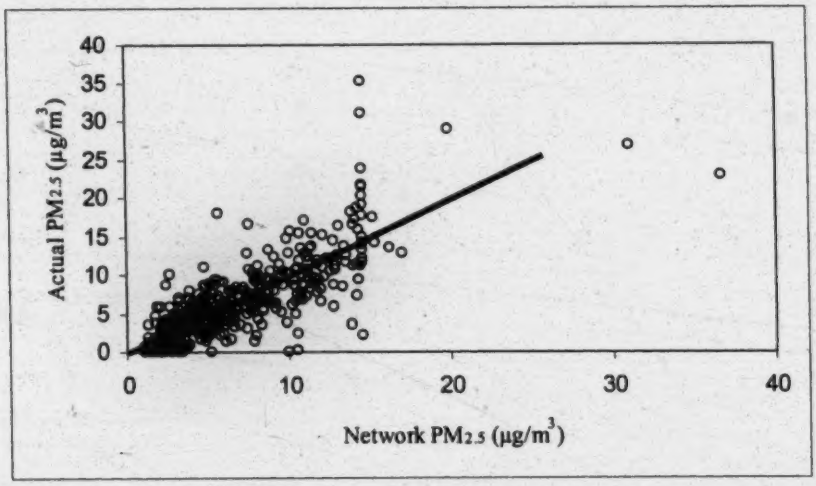


Figure B- 25 Network vs. Actual $PM_{2.5}$ for Calgary East spring forecast model with lagged data as an input.

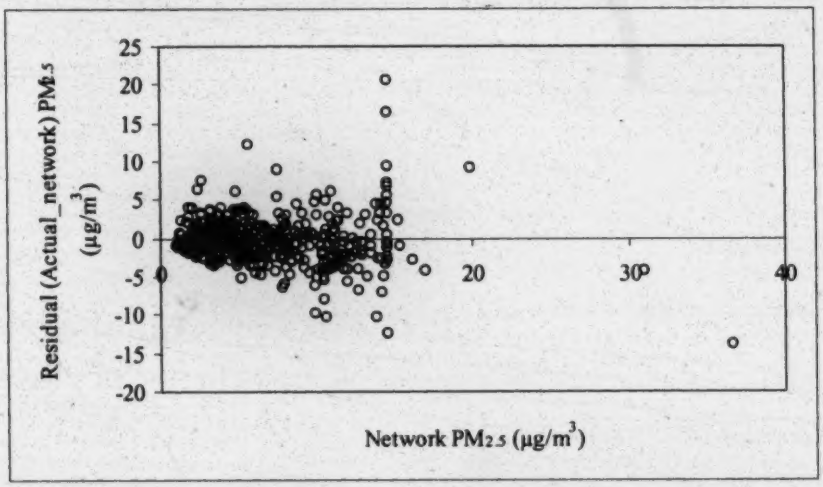


Figure B- 26 Calgary East spring forecast model (lagged data) residual analysis.

Summer season

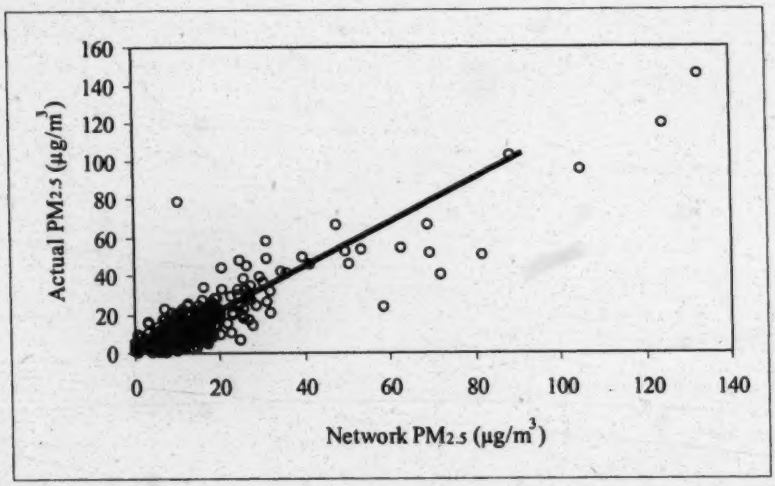


Figure B- 27 Network vs. Actual $PM_{2.5}$ for Calgary East summer forecast model with lagged data as an input.

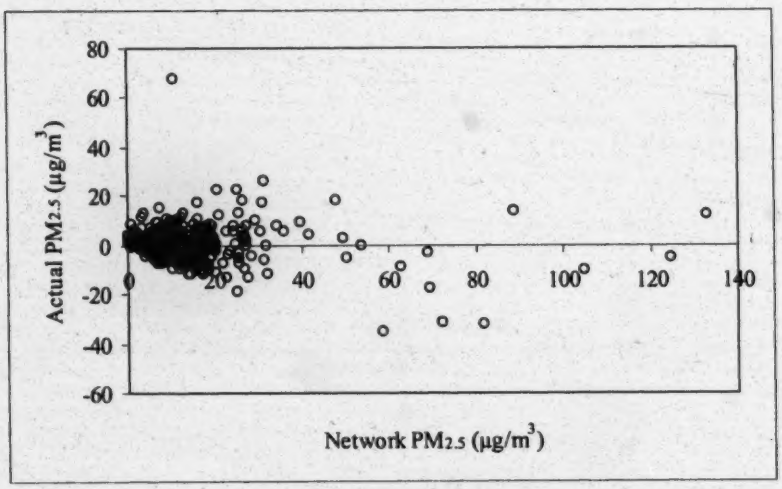


Figure B- 28 Calgary East summer forecast model (lagged data) residual analysis

Fall season

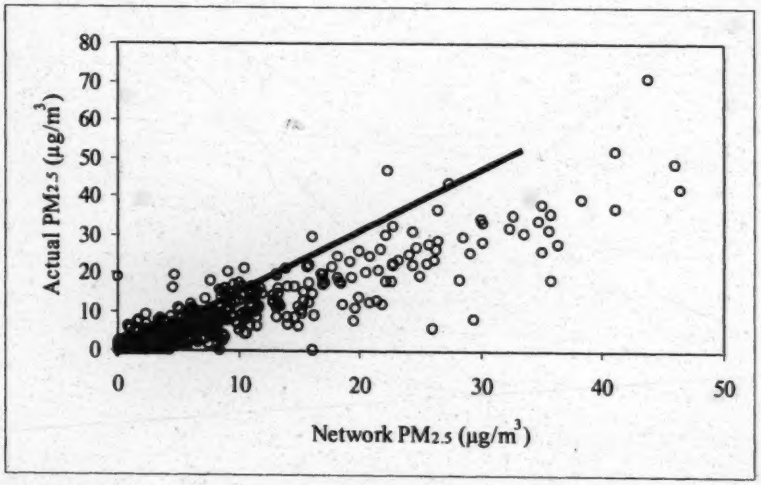


Figure B- 29 Network vs. Actual $PM_{2.5}$ for Calgary East fall forecast model with lagged data as an input.

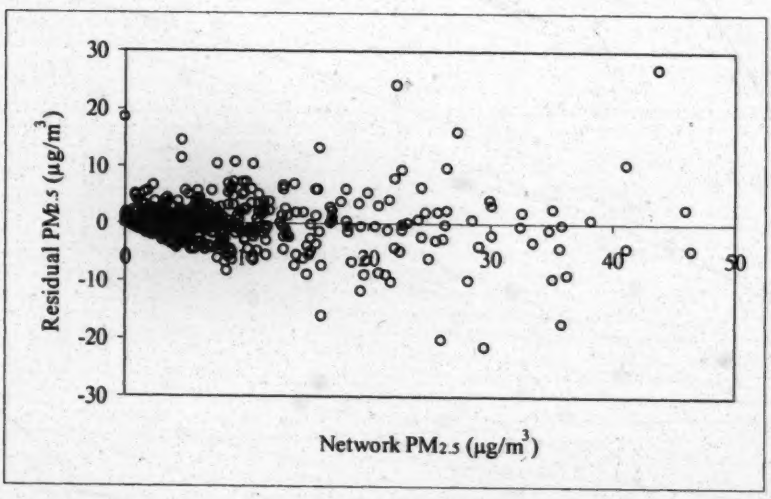


Figure B- 30 Calgary East fall forecast model (lagged data) residual analysis

Winter season

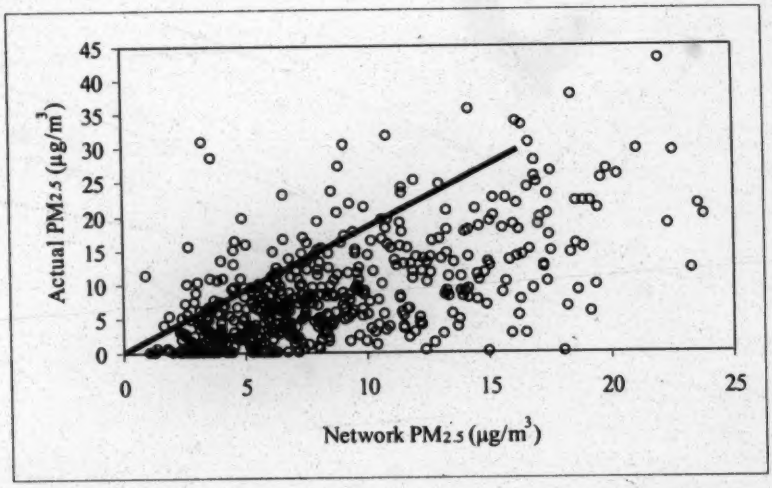


Figure B- 31 Network vs. Actual $PM_{2.5}$ for Calgary East winter forecast model with lagged data as an input.

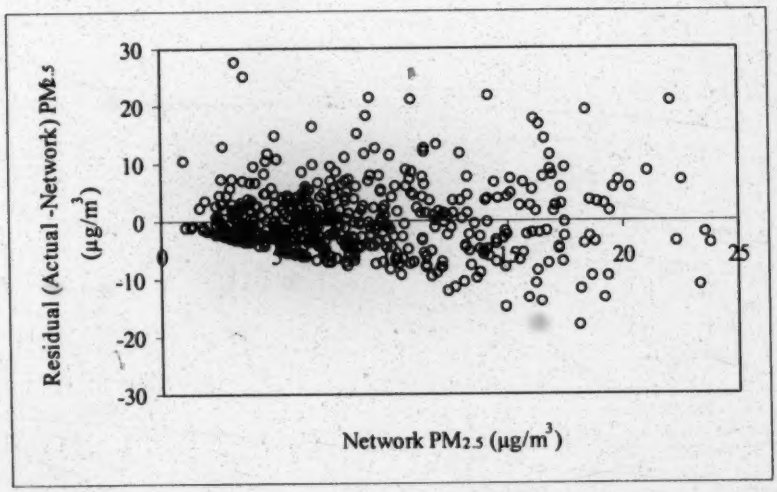


Figure B- 32 Calgary East winter forecast model (lagged data) residual analysis